

Montpellier SupAgro

THESE

Pour obtenir le grade de

DOCTEUR EN SCIENCES DE MONTPELLIER SupAgro

Discipline : Biologie Intégrative des Plantes

Ecole Doctorale : SIBAGHE

Préparée au CIRAD – Département BIOS / UMR DAP / Equipe SRG

Présentée et soutenue publiquement

par

Daniel Foncéka

Le 10 décembre 2010

Titre : Elargissement de la base génétique de l'arachide cultivée (*Arachis hypogaea*) : Applications pour la construction de populations, l'identification de QTL et l'amélioration de l'espèce cultivée

JURY

A. CHARRIER Professeur émérite, Montpellier SupAgro
G. CHARMET Directeur de recherche, INRA
L. GRIVET Directeur de recherche, Syngenta
F. HOSPITAL Directeur de recherche, INRA
B. COURTOIS Chercheur, CIRAD
J.-F. RAMI Chercheur, CIRAD

Président
Examineur
Rapporteur
Rapporteur
Directeur de Thèse
Co-encadrant

RESUME

L'arachide (*Arachis hypogaea* L.) est une plante allotétraploïde ($2n = 4x = 40$) issue d'un événement récent d'hybridation entre deux espèces sauvages suivi d'un doublement spontané des chromosomes. La variabilité génétique existant dans le compartiment cultivé est faible. Les espèces diploïdes sauvages apparentées à l'arachide cultivée représentent un important réservoir d'allèles nouveaux utilisables pour élargir le pool génique du cultigène. L'objectif général de cette thèse est d'augmenter les marges de progression en amélioration de l'arachide par recours aux ressources génétiques sauvages.

Nous avons développé une population BC_1F_1 à partir du croisement entre Fleur11 et un amphidiploïde synthétique (*A. ipaensis* x *A. duranensis*)^{x4} qui associe les génomes des plus probables progéniteurs sauvages de l'espèce cultivée. Nous avons d'abord construit une carte génétique comprenant 298 loci positionnés sur 21 groupes de liaison avec une taille totale de 1843.7 cM. Nous avons ensuite conduit une analyse AB-QTL pour plusieurs caractères impliqués dans la productivité, l'adaptation et la domestication de l'arachide. Au total, nous avons cartographié 95 QTL. Pour plusieurs QTL, les effets positifs sont associés aux allèles provenant des espèces sauvages. Nous avons aussi identifié trois régions du génome qui portent des empreintes de la domestication. Nous avons enfin développé une population de 122 lignées d'introgression à l'aide d'une stratégie de sélection assistée par marqueurs. L'ensemble des groupes de liaison sont couverts avec des fragments chevauchants, issus des donneurs sauvages, d'une taille moyenne de 39.2 cM et chaque lignée comprend en moyenne 1.2 fragments. Nous avons par ailleurs discuté l'utilisation de cette collection de lignées d'introgression pour des applications de sélection et de génétique. Nos résultats ouvrent de nouvelles perspectives pour l'amélioration de l'arachide par croisement avec les espèces sauvages apparentées.

Mots-clés : *Arachis hypogaea*, cartographie de QTL, domestication, ressources génétiques sauvages, polyploïdes, tolérance à la sécheresse, amélioration.

ABSTRACT

Peanut (*Arachis hypogaea* L.) is considered to be an allotetraploid ($2n = 4x = 40$) originated from a single hybridization event between two wild diploids followed by spontaneous chromosomes duplication. Cultivated peanut harbours a limited genetic diversity. Peanut wild relatives represent an important source of novel alleles that could be used to broaden the gene pool of the cultigen. The general objective of this work is to enhance the rate of progress in peanut breeding using wild species resources.

We developed a BC_1F_1 population from the cross between Fleur11 and a synthetic amphidiploid (*A. ipaensis* x *A. duranensis*)^{x4} that combines the genomes of the most probable wild progenitors of the cultigen. We first developed a SSR based genetic map of 298 loci on 21 linkage groups, spanning a total map distance of 1843.7 cM. We then conducted a detailed AB-QTL analysis for several traits involved in peanut productivity and adaptation as well as in the domestication syndrome. We mapped a total of 95 QTLs. About half of the QTL positive effects were associated with alleles of the wild parents. We identified three genomic regions which bear footprints of domestication. We finally developed, through a marker assisted backcross strategy, an exotic introgression library of 122 lines. This population, which has in average 1.2 fragments per line, allows covering all linkage groups with overlapping wild donor fragments of 39.2 cM length in average.

The utilization of the exotic introgression lines library for breeding and genetics is discussed. Our findings open new avenues for peanut improvement using wild relatives.

Keywords: *Arachis hypogaea*, QTL mapping, domestication, wild relatives, polyploid, drought tolerance, breeding.

REMERCIEMENTS

Ah ! Nous y voilà ! Je commençais à ne plus y croire...

Des études pifométriques ont montré que les remerciements sont très certainement la rubrique de la thèse la plus « widely consulted ». Alors or-ga-ni-sons nous ! Une structuration en chapitres s'impose.

Chapitre I. Les officiels

Mes remerciements s'adressent aux institutions CIRAD, GCP et CERAAS pour avoir créé les conditions financière et d'accueil nécessaires à la réalisation de ces travaux de thèse. Je voudrais remercier tout particulièrement Jean-François Rami et Jean-Christophe Glaszmann pour avoir su mobiliser les ressources financières durant ces trois années de thèse et Ousmane Ndoeye pour avoir facilité mon accueil au CERAAS et au CNRA de Bambey durant mes séjours au Sénégal. J'aimerais aussi remercier chaleureusement ma directrice de thèse, Brigitte Courtois, pour sa rigueur et ses encouragements, Jean-François Rami, pour son encadrement de proximité, et les membres de mon comité de thèse : Mathilde Causse, Mathias Lorieux, David Bertioli, Nourollah Ahmadi, Angélique D'Hont et Delphine Luquet pour leurs conseils avisés. Mes remerciements vont aussi à l'endroit des scientifiques émérites (c'est fayot, je sais !) qui ont bien voulu participer à mon Jury de thèse : André Charrier, Gilles Charmet, Frédéric Hospital, Laurent Grivet, Brigitte Courtois et Jean-François Rami (encore lui ! et ce n'est pas fini, il s'est investi il sera remercié...).

Mes remerciements s'adressent aussi à tous les chercheurs de l'équipe SRG.

Chapitre II. Les collègues et amis

I. 1 Du CIRAD et du CERAAS

Si le doctorat était un diplôme communautaire, je l'aurai partagé volontiers avec Tossim, Ronan, Hélène, Elodie et Fabien car sans leur disponibilité et leur professionnalisme je ne serai très certainement pas arrivé au bout de cette thèse. Cependant, la thèse étant une production d'UN esprit, l'aboutissement d'UN travail de recherche, UN exercice académique validé par l'obtention d'UN grade universitaire... (lu dans le guide pour la rédaction et la présentation des thèses), je ne peux donc pas en partager la sanction avec vous. Je me contenterai de vous remercier.

Je voudrais aussi remercier mes collègues doctorants d'hier, Sophie, Julie et d'aujourd'hui, Julien, avec qui j'ai partagé, en plus d'un bureau, mes réflexions prospectives et souvenirs d'Ancien, problèmes de papiers et visions politiques...

Merci à Claire Billot (source inépuisable d'info pratiques et d'astuces) et à David Pot (Grand Emprunteur devant l'éternel !) pour leur présence et les moments de rigolade (et il y en a eu !) durant ces trois années.

J'adresse aussi mes remerciements à Aida pour sa générosité, son hospitalité et sa joie de vivre, à Mbaye Ndoeye Sall, Issa Faye et aux compagnons doctorants du Ceraas : Roger (roi du agouti fumé), Ambroise (ambassadeur du kédjénou), Sali, Patrice, Eric, que la bière soit avec vous ! Un grand merci à la famille Muller sise à « Angueule Meule » Thiès.

Merci à Angel (pourquoi pas nous ?), à Céline (rdv dans 10 ans !), Cécile (non je ne joue pas dans Dr House !), Mr Mournet (directeur sportif rugby touch), à Jean-Baptiste (lo moughn mou diekh !), Caroline (je ne sais pas quoi dire), Delphine (tu tiens le bon bout), Olivier et Cyril (vive St Malo et la polyploïdie), Mumu, Giang, Camille, Ablanvi...

I. 2 D'ici et d'ailleurs

Merci à toute la bande de joyeux lurons avec qui j'ai passé de supers moments à Montpellier : Baptiste Adam, Hélène Vergne, Grogro Barnaud, Adeline Bottalico, Eric Mirouz, Marie Lacombe, Jean-François Pak, Lai Ting Rami, Patrick Berthouly, Anne Heuret...et au sponsor officiel de nos fêtes : la brasserie des Garrigues (elle est quand même bonne hein !)

Chapitre III. Les familles

J'aimerais dire un grand merci à mes parents qui ont toujours pensé que je ferais de longues études. Ils me disaient : « tu me feras de grandes études mon fils ». J'en étais pas particulièrement persuadé, je ne sais pas non plus d'où ils tenaient ces infos et cette confiance en moi, mais ça a dû très certainement me motiver à toujours donner le meilleur de moi-même. J'aimerais aussi remercier mes frères Loulou, Yvon et Jean-Michel pour leur soutien sans faille (en oulof ça se dit « dem ba diekh ») en toute circonstance, mes sœurs, mes beaux-frères, mes neveux et nièces (je vous passe les prénoms parce que « famille africaine là c'est beaucoup dé ! ») et enfin tonton Dief (un oncle d'emprunt).

Un grand merci à ma belle famille Dominique, Rosalie, Aymeric et Marie pour leur soutien.

Chapitre IV. Les saugrenus

Merci aux génies créateurs Ramiesque et Danielesque, pour avoir conçu le polyapplicateur®, structure en PVC plate et percée d'orifices régulièrement répartis, permettant de plonger simultanément 38 tubes eppendorf (2 ml) dans un récipient d'azote liquide et de les disposer (toujours avec la même simultanéité !) dans un bloc pour vibro-broyeur MM301 sans se brûler les doigts !

Merci au feuilleton politico-judiciaire Woerth-Bettencourt qui a permis de combler, par l'abondance de ses rebondissements, toutes mes périodes de panne d'inspiration.

Chapitre V. Les hors-catégories

J'aimerais enfin dire un immense merci à ma femme Aurélie qui m'a toujours soutenu dans toutes mes entreprises, mais particulièrement pendant ces trois années de thèse. Tu as su supporter et gérer mes absences, mes périodes stress, mes angoisses, trouver les mots justes pour me motiver dans les moments de doute. Merci encore pour ta confiance et ta patience.

Merci à ma fille Telma, qui n'a pas toujours compris mes longues absences pour mission au Sénégal mais qui a su les accepter avec une maturité surprenante pour un enfant de cet âge !

Merci enfin à « Yves », le petit dernier né cette année pour tout le bonheur que tu nous apportes.

TABLE DES MATIERES

AVANT-PROPOS	6
OBJECTIFS ET PRESENTATION DE LA THESE	9
CHAPITRE I ANALYSES BIBLIOGRAPHIQUES.....	10
I. GENERALITES SUR L'ARACHIDE.....	10
I-1 <i>Aspects socioeconomiques.....</i>	<i>10</i>
I-1-1 Production mondiale et importance economique	10
I-1-2 Utilisation, importance nutritionnelle et risques associés	12
I-2 <i>Origine et distribution.....</i>	<i>14</i>
I-3 <i>Biologie et mode de reproduction de l'espèce cultivée</i>	<i>14</i>
I-4 <i>Systématique et taxonomie.....</i>	<i>15</i>
I-5 <i>Description des espèces de la section Arachis.....</i>	<i>16</i>
I-6 <i>Relation entre espèces diploïdes de la section Arachis.....</i>	<i>17</i>
I-7 <i>Les espèces diploïdes sauvages à l'origine de l'espece cultivée.....</i>	<i>18</i>
I-8 <i>Diversité génétique de l'espèce cultivée</i>	<i>20</i>
I-9 <i>Amélioration génétique de l'arachide cultivée</i>	<i>21</i>
I-9-1 Amélioration de la tolérance à la sécheresse	23
I-9-2 Amélioration de la résistance aux maladies.....	26
I-9-3 Amélioration de la qualité des huiles et de la qualité gustative	28
II. UTILISATION DES RESSOURCES GENETIQUES SAUVAGES POUR L'AMELIORATION VARIETALE.....	30
II-1 <i>Domestication et réduction de la variabilité génétique des espèces cultivées</i>	<i>30</i>
II-2 <i>La détection de QTL a partir de croisements interspécifiques et les applications pour la création variétale.....</i>	<i>32</i>
II-2-1 la cartographie de QTL dans des populations de backcross avancés (AB-QTL)	32
II-2-2 Les lignées d'introgression	35
II-2-2-1 Potentiel pour la détection et la validation de QTL	35
II-2-2-2 Potentiel pour la création variétale assistée par marqueur	39
CHAPITRE II GENETIC MAPPING OF WILD INTROGRESSIONS INTO CULTIVATED PEANUT: A WAY TOWARD ENLARGING THE GENETIC BASIS OF A RECENT ALLOTETRAPLOID	41
CHAPITRE III FOSTERED AND LEFT BEHIND ALLELES IN PEANUT (<i>ARACHIS HYPOGAEA</i> L.): INTERSPECIFIC QTL MAPPING REVEALS FOOTPRINTS OF DOMESTICATION AND USEFUL NATURAL VARIATION FOR BREEDING.....	55
CHAPITRE IV CONSTRUCTION D'UNE COLLECTION DE LIGNEES D'INTROGRESSION A PARTIR D'UN CROISEMENT INTERSPECIFIQUE CHEZ L'ARACHIDE.....	86
I. BREF RAPPEL DE LA PROBLEMATIQUE.....	86
II. MATERIEL ET METHODES	87
III. RESULTATS ET DISCUSSION.....	93
III-1 <i>Choix de la population d'étude.....</i>	<i>93</i>
III-2 <i>Développement de la collection de lignées d'introgression.....</i>	<i>94</i>
IV. CONCLUSION	96
CHAPITRE V CONCLUSION GENERALE ET PERSPECTIVES.....	97
I. LA SATURATION DE LA CARTE GENETIQUE.....	100
II. CARACTERISATION DE LA POPULATION DE LIGNEES D'INTROGRESSION.....	101
III. CREATION VARIETALE PAR PYRAMIDAGE DE QTL ASSISTEE PAR MARQUEURS.....	103
IV. ORIGINE ET DOMESTICATION DE L'ESPECE CULTIVEE.....	104
V. CARTOGRAPHIE FINE ET CLONAGE DES GENES SOUS-JACENTS AUX QTL.....	105
VI. VERS LA CONSTRUCTION DE LARGES COLLECTIONS DE LIGNEES D'INTROGRESSION	106
REFERENCES.....	108

AVANT-PROPOS

Ce travail de thèse s'inscrit dans une dynamique de recherche internationale sur l'arachide soutenue par des consortiums de recherche et le CIRAD. Il a bénéficié d'un contexte scientifique et partenarial favorable, tant d'un point de vue international avec l'échange de matériel génétique entre le Brésil, la France et le Sénégal et l'augmentation constante des ressources génomiques, que national avec l'appui technique du programme de sélection arachide du CNRA de Bambey (ISRA, Sénégal) et logistique du CERAAS (ISRA). Dans cet avant-propos, nous faisons un bref historique des projets successifs ayant permis de réaliser ce travail et positionnons la thèse dans le contexte du programme de sélection arachide de l'ISRA et de l'équipe SRG/EGV de l'UMR DAP du CIRAD. .

Au cours de cette dernière décennie, la communauté scientifique, par le biais du consortium international Generation Challenge Program (GCP) (www.generationcp.org), a investi d'importants efforts de recherche sur les espèces cultivées parmi lesquelles quelques plantes dites « orphelines » en termes de ressources génomiques. Ces recherches ont permis, au travers d'un partenariat multisectoriel et interdisciplinaire, d'accroître les connaissances sur la diversité génétique de plantes cultivées et de développer des outils génétiques et génomiques pour une meilleure utilisation des ressources génétiques dans l'amélioration des plantes. En plus de l'accroissement et de la mise en commun des connaissances, des ressources génétiques et des outils, le consortium a favorisé le développement de partenariats originaux impliquant les Centres Internationaux de Recherche (CGIAR), des Instituts Avancés de Recherche, des Universités et des Systèmes Nationaux de Recherche Agricoles (SNRA). Dans ce contexte, le CIRAD (France), l'ISRA (Sénégal), l'EMBRAPA (Brésil), et l'ICRISAT (Inde) sont, depuis 2004, partenaires dans des projets de recherche sur l'utilisation des espèces diploïdes sauvages apparentées à l'arachide pour l'élargissement de la base génétique et l'amélioration de l'espèce cultivée. C'est ainsi qu'en 2004, dans un premier projet "Unlocking the genetic diversity in peanut's wild relatives with genomic and genetic tools", des tétraploïdes artificiels ou amphidiploïdes, ayant le même niveau de ploïdie que l'arachide cultivée, impliquant des génomes diploïdes sauvages A et B ont été développés par l'EMBRAPA. Ces amphidiploïdes ont été utilisés par le programme de sélection arachide de l'ISRA pour produire plusieurs populations de backcross à partir de croisements interspécifiques. De plus, ce projet a permis de développer des banques d'EST et des banques BAC, d'augmenter considérablement le nombre de marqueurs moléculaires rendant ainsi

possible la production de cartes génétiques et l'application des méthodes modernes de sélection chez l'arachide. Une des populations de backcross interspécifiques produite par ce projet a été utilisée comme matériel de départ dans ce travail de thèse. Le financement des activités de recherche de la thèse a été supporté en partie par deux nouveaux projets intitulés "Application of molecular tools for controlled wild introgression into peanut cultivated germplasm in Senegal" et "Enhancing groundnut (*Arachis hypogaea* L.) genetic diversity and speeding its utilization in breeding for improving drought tolerance" financés en 2007 et 2008 par le GCP. En parallèle, depuis 2006, la fondation Bill et Melinda Gates soutient un important programme de recherche sur les légumineuses tropicales. Ces financements ont largement profité à la communauté de recherche sur l'arachide et ont permis d'accroître les connaissances sur la diversité génétique de l'arachide et sa structuration, sur la synténie avec des espèces modèles (*Medicago*, *Lotus*), d'augmenter le nombre de marqueurs moléculaires, le nombre de populations de sélection, et de développer de nouveaux dispositifs pour le phénotypage.

Comme nous l'avons indiqué plus haut, cette thèse a été menée en étroite collaboration avec le programme de sélection arachidière de l'ISRA. Ce programme de sélection a une longue expérience dans l'amélioration génétique de l'arachide et compte déjà plusieurs obtentions variétales à son actif. Les principales cibles de sélection du programme sont l'amélioration pour la tolérance à la sécheresse, la résistance aux cercosporioses et la dormance des graines fraîches. L'utilisation des espèces sauvages comme source de caractères d'intérêts agronomiques pour l'élargissement de la base génétique et l'amélioration de l'arachide cultivée est une impulsion nouvelle donnée au programme de sélection arachidière de l'ISRA et il s'est développé, au cours des projets successifs, une solide expertise dans les domaines de l'hybridation interspécifique et de la conduite de populations en serre et en champ. Dans cette thèse, toutes les activités de construction de populations ont été exécutées au CERAAS et les activités de phénotypage au champ au CNRA de Bambey.

Au cours de cette dernière décennie, on a assisté à une rapide évolution des unités et équipes de recherche impliquées dans la génomique et l'amélioration des plantes au CIRAD. La modification de la structuration des unités de recherche a aussi entraîné une évolution des projets de recherche des différentes équipes. L'équipe Structuration des Ressources Génétiques (SRG), au sein de laquelle cette thèse a été effectuée, a récemment évolué vers un nouveau projet scientifique qui porte sur l'Evaluation la Gestion et la Valorisation des ressources génétiques (EGV). Les objectifs de ce nouveau projet portent sur la définition de stratégies d'analyse de la diversité génétique, l'identification des facteurs biologiques,

environnementaux et sociaux qui ont conduit à la structuration de cette diversité et, la valorisation des connaissances sur les ressources génétiques dans les programmes de sélection. L'arachide est une des espèces qui a été retenue dans le dispositif de recherche de l'équipe parce qu'elle offre des perspectives de recherche intéressantes en termes d'identification des facteurs génétiques liés à la domestication de l'espèce cultivée, d'analyse des relations entre les espèces sauvages et l'espèce cultivée et de développement de schémas de sélection innovants pour une utilisation des ressources génétiques sauvages dans les programmes d'amélioration de l'espèce cultivée. Les activités de recherche développées au cours de cette thèse entrent dans le cadre du nouveau projet de recherche de l'équipe.

OBJECTIFS ET PRESENTATION DE LA THESE

L'objectif principal de ce travail de thèse porte sur l'élargissement de la base génétique de l'arachide cultivée par l'utilisation de ressources génétiques sauvages. Plus spécifiquement, cette thèse s'est focalisée sur l'identification d'allèles d'origine sauvage utilisables en sélection pour l'amélioration de caractères d'intérêt agronomique, sur l'identification de régions génomiques qui ont été impliquées dans la domestication de l'espèce cultivée et sur la construction d'une population de lignées d'introgression.

Cette thèse s'articule autour de cinq chapitres. Le premier chapitre fait l'état des lieux des connaissances sur l'arachide et les espèces sauvages apparentées. Il aborde aussi, à l'aide d'exemples sur plusieurs espèces, les aspects relatifs aux stratégies mises en place pour l'utilisation des espèces sauvages dans les programmes d'amélioration génétique basés sur des croisements interspécifiques.

Le second chapitre, qui a fait l'objet d'une publication parue dans la revue BMC Plant Biology, porte sur la construction de la carte génétique à partir d'une population BC₁F₁ issue du croisement entre la variété cultivée Fleur11 et l'amphidiploïde sauvage (*A. ipaensis* x *A. duranensis*)^{x4}. Il aborde aussi l'analyse de la colinéarité entre les génomes A et B et l'étude de l'évolution de la taille des fragments d'ADN d'origine sauvage entre la population BC₁F₁ et la population BC₂F₁.

Le troisième chapitre, qui a aussi fait l'objet d'une publication soumise à la revue Genetics, présente les résultats de l'analyse AB-QTL. Dans ce chapitre sont abordés la détection de QTL sous deux régimes hydriques, l'identification de zones du génome qui portent des empreintes de la domestication et l'analyse de la distribution des QTL entre les génomes A et B.

Dans le quatrième chapitre sont décrites les différentes stratégies que nous avons utilisées pour la construction d'une population de lignées d'introgression.

Le cinquième chapitre est une synthèse des résultats majeurs obtenus au cours de cette thèse. Il ouvre des perspectives de recherche en termes d'utilisation des lignées d'introgression pour la validation et la cartographie fine des QTL ainsi que pour le clonage des gènes sous-jacents aux QTL.

Chapitre I ANALYSES BIBLIOGRAPHIQUES

I. GENERALITES SUR L'ARACHIDE

I-1 ASPECTS SOCIOECONOMIQUES

I-1-1 PRODUCTION MONDIALE ET IMPORTANCE ECONOMIQUE

L'arachide, douzième production végétale dans le monde, est une culture majeure dans la plupart des régions tropicales et subtropicales. Elle est cultivée sur tous les continents, dans 120 pays environ, sur une superficie totale de 24.6 millions d'hectares pour une production de 38.2 millions de tonnes (FAOSTAT, 2008). Les pays les plus grands producteurs d'arachide se retrouvent sur les continents asiatique, africain et américain. Le continent asiatique avec 13.3 millions d'hectares regroupe la moitié des superficies mondiales emblavées en arachide. La Chine (13 millions de tonnes) et l'Inde (9 millions de tonnes) sont les premiers producteurs mondiaux avec une production qui a contribué pour plus de la moitié de la production mondiale en 2007. Les USA (2.3 millions de tonnes) sont le premier pays producteur du continent américain. Le continent africain, avec ses 10 millions d'hectares de surfaces occupées par la culture de l'arachide et ses 10 millions de tonnes occupe la seconde place devant le continent américain (FAOSTAT, 2008). La production d'arachide du continent africain a connu une croissance importante depuis le début des années 1990. Cette croissance est principalement liée à l'augmentation de la production dans les pays d'Afrique de l'Ouest (Revoreda and Fletcher, 2002). Par exemple le Nigéria, troisième producteur mondial, a contribué pour plus du tiers de la production arachidière du continent africain en 2008. Il devance, dans ce continent, le Soudan (716.000 tonnes) et le Sénégal (647.000 tonnes). La tendance à l'augmentation de la production arachidière est observée dans la plupart des grands pays producteurs à l'exception du Sénégal. Dans ce pays, une réduction de 25.8% de la production a été observée en comparant la moyenne de la production des années 1972 à 1975 à celle des années 1996 à 2000 (Revoreda and Fletcher, 2002). La baisse de la production arachidière au Sénégal résulte principalement des perturbations climatiques et des déficiences d'approvisionnement en intrants, particulièrement en semences de qualité (Lettre de politique de développement de la filière arachide, Gouvernement du Sénégal, 2003).

L'Afrique, bien que deuxième continent en termes de production d'arachide, possède les rendements à l'hectare les plus faibles avec en moyenne 1 t/ha, comparé à l'Asie (1.8 t/ha) et

à l'Amérique (3 t/ha). De manière générale depuis les années 1970, on a assisté à une augmentation tendancielle des rendements dans les trois grandes régions de production de l'arachide (Asie, Afrique et Amérique). La plus forte progression des rendements a été observée en Asie : 74.9% lorsque l'on compare la moyenne des rendements des années 1972 à 1975 à celle des années 1996 à 2000. Pour les mêmes périodes, la progression est de 24.9% pour l'Amérique et de 15.2% pour l'Afrique (Revoredo and Fletcher, 2002). Les rendements dans les grands pays producteurs varient d'un peu plus de 1.5 t/ha (Inde, Nigéria) à un peu plus de 3 t/ha (Chine, USA) et peuvent atteindre plus de 6 t/ha dans certains pays comme Israël (FAOSTAT, 2008). Dans la plupart des pays d'Afrique au sud du Sahara, les rendements sont inférieurs à 1 t/ha. Ces faibles rendements sont liés au système de culture pluvial associé à de très faibles apports d'intrants. Dans ces pays, soumis aux aléas et à l'irrégularité des événements pluvieux pendant la saison culturale, on note aussi une forte variation interannuelle des rendements. Au Sénégal par exemple, on a observé une forte fluctuation des rendements entre 1998 et 2008, avec des pics de plus de 1 t/ha en 1998 et 1999 et des creux en 2003 et 2007 avec 0.3 t/ha et 0.5 t/ha, respectivement.

L'arachide représente environ 10% de la production mondiale d'oléagineux (Fletcher and Nadolnyak, 2006). Son volume de production a représenté un chiffre d'affaires d'environ 17 milliards de dollars en 2007 (FAOSTAT). Les exportations d'arachide ne représentent que 6% de la production mondiale car la majeure partie est consommée localement. Les premiers pays exportateurs d'arachide au monde sont la Chine, les USA, l'Argentine, l'Inde et le Vietnam. Les trois premiers pays cités contribuent pour plus de 70% des exportations mondiales (Boonsaeng and Fletcher, 2008). Ces exportations se font majoritairement en direction des pays de l'Union Européenne, du Japon, du Mexique, de la Russie, de l'Indonésie et du Canada. Le volume des exportations vers les pays de l'Union Européenne n'a cessé de croître depuis la fin des années 90 passant de 612 000 tonnes en 1999 à environ 758 000 tonnes pour une valeur totale d'environ 527 millions d'Euros en 2005 (Boonsaeng and Fletcher, 2008). L'arachide est l'une des principales cultures de rente dans les pays en voie de développement, plus particulièrement en Afrique subsaharienne où l'agriculture contribue pour 30 à 50% du PIB. La culture de l'arachide a constitué une grande partie des exportations de certains pays d'Afrique de l'Ouest. Jusque dans les années 1970, les pays africains représentaient les premiers exportateurs d'arachide au monde. Au Sénégal par exemple, l'arachide a compté pour 80% des recettes d'exportations dans les années 1960. L'évolution des marchés mondiaux vers une demande d'arachide à usage alimentaire a fait chuter les volumes d'exportation des pays africains, d'autant plus que leur production était affectée par

la présence d'aflatoxine. Aujourd'hui, au Sénégal par exemple, la part qu'occupe le volume des exportations dans le PIB a beaucoup diminué mais la culture de l'arachide reste la source de revenu principale pour les petits producteurs.

I-1-2 UTILISATION, IMPORTANCE NUTRITIONNELLE ET RISQUES ASSOCIES

L'arachide produite dans le monde est principalement transformée en huile, en farine et en dérivés qui entrent dans la composition de produits alimentaires autres (confiserie, beurre de cacahuète, pâte d'arachide...). Durant la période 1996-2000, 49.2% de la production mondiale a servi à produire de l'huile et de la farine et 41.1% est entrée dans la composition de produits alimentaires (Revoredo and Fletcher, 2002). Les parts qu'occupent les différentes formes d'utilisation de l'arachide sont variables en fonction des régions et à l'intérieur des régions en fonction des pays. En Asie par exemple, la Chine consacre environ 40% de sa production à des fins alimentaires contre seulement 9% en l'Inde. La part de la production arachidière destinée à l'alimentation est de 77% aux Etats-Unis et de 55.3% en Afrique de l'Ouest. Dans cette dernière région, la part de la production arachidière destinée à l'alimentation a augmenté de 209% en comparant les moyennes des années 1972-1975 à celles des années 1996-2000 (Revoredo and Fletcher, 2002).

L'arachide est aussi utilisée pour l'alimentation animale. Par exemple, le tourteau qui résulte de l'extraction d'huile constitue un aliment riche en protéines pour le bétail. De même, dans la plupart des pays du Sahel, les fanes sont séchées après la récolte et apportées en foin pour l'alimentation animale. Elles constituent souvent l'un des principaux apports alimentaires pour les petits ruminants pendant la saison sèche.

Récemment, la montée des cours mondiaux du pétrole a provoqué une forte demande en énergies alternatives non-fossiles. La filière oléagineuse a été très tôt identifiée pour servir de biocarburant. Dans cette filière, l'huile d'arachide est de plus en plus présentée comme source potentielle pour la production de biocarburant. Cependant les coûts liés à sa production et la demande importante en huile alimentaire rendent l'huile d'arachide moins compétitive que les autres huiles d'origine végétale comme le soja par exemple. Pour pallier les problèmes de compétitivité de l'huile d'arachide, des recherches sont en cours à l'Université de Géorgie pour produire des variétés d'arachide à forte teneur en huile, impropres à la consommation humaine et à faible besoin en intrants. La variété « Georganic » qui semble répondre à tous ces critères est présentée comme l'avenir des biocarburants à base d'huile d'arachide (http://southeastfarmpress.com/mag/farming_peanut_biodiesel_promising).

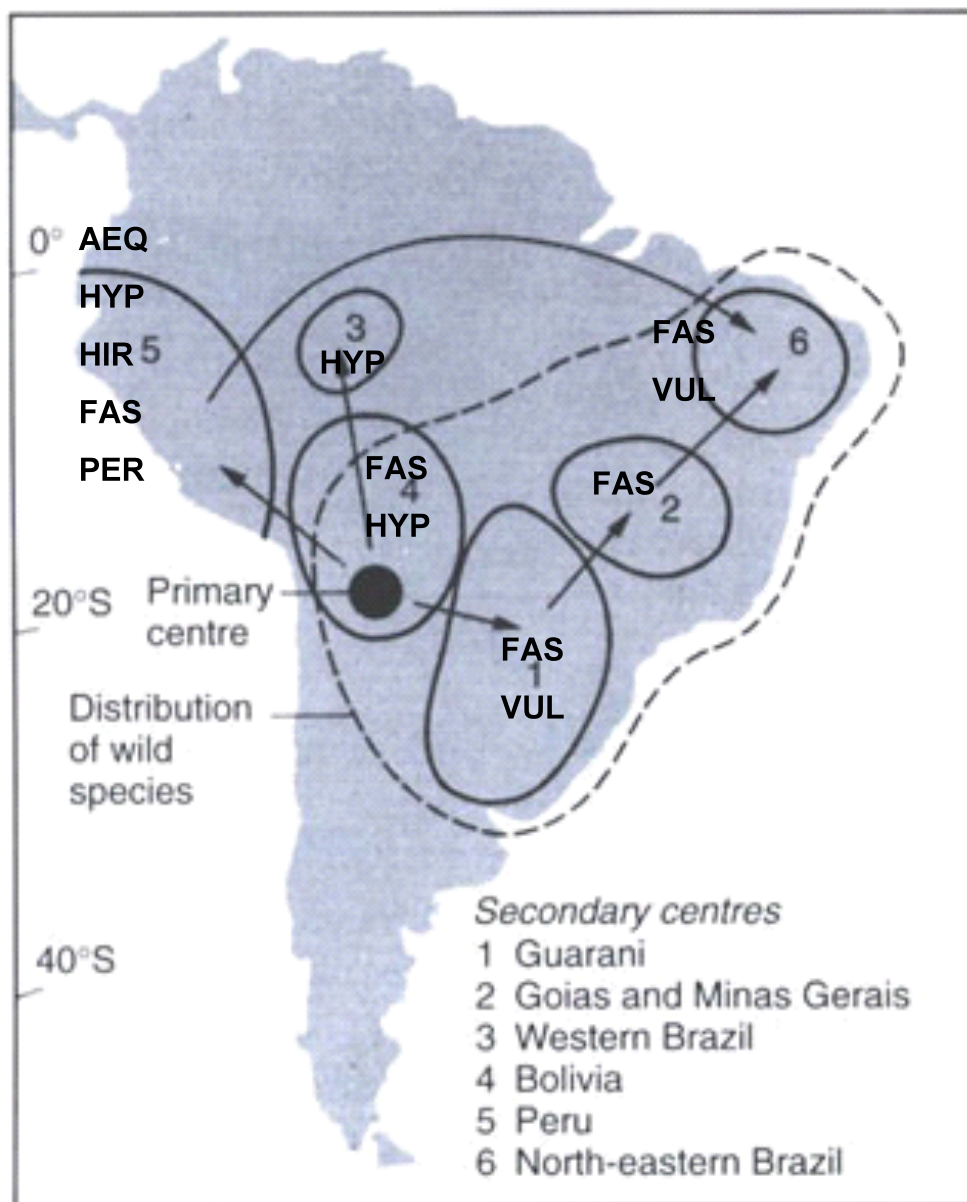


Figure I.1 : Centre d'origine et distribution de l'arachide cultivée. (Figure adaptée de Ferguson et al., 2004). **HYP** : *A. hypogaea ssp. hypogaea var. hypogaea*, **FAS** : *A. hypogaea ssp. fastigiata var. fastigiata*, **VUL** : *A. hypogaea ssp. fastigiata var. vulgaris*, **HIR** : *A. hypogaea ssp. hypogaea var. hirsuta*, **PER** : *A. hypogaea ssp. fastigiata var. peruviana*, **AEQ** : *A. hypogaea ssp. fastigiata var. aequatoriana*.

L'arachide, comme la plupart des plantes légumineuses à graines, occupe une place importante dans l'alimentation humaine. C'est une bonne source de lipides, de protéines et de sels minéraux. Les graines contiennent environ 45-50% de lipides, 25-30% de protéines, 5-12% de carbohydrates et 3% de fibres. Des études menées aux Etats-Unis ont montré qu'une consommation bihebdomadaire d'arachide et/ou de produits dérivés améliorait la qualité des régimes alimentaires (Griel et al., 2004). Cette amélioration se caractérisait par une augmentation des apports en protéines, en matières grasses totales, en acides gras poly et mono-insaturés, en fibres, en vitamine E, en acide folique, en sels minéraux (calcium, magnésium, zinc et fer) et par une augmentation globale des apports énergétiques. Plusieurs autres études ont aussi montré l'absence d'association entre la consommation d'arachide et la tendance à la prise de poids (Hu et al., 1998; Sabate, 2003). Par ailleurs des études médicales ont montré que la consommation de noix en général et d'arachide en particulier réduisait les risques de maladies cardiovasculaires (Fraser, 2000; Albert et al., 2002). Les valeurs nutritives de l'arachide ont été récemment mises à profit dans la composition d'aliments à haute valeur nutritive utilisés pour le traitement de la malnutrition sévère chez l'enfant (Briend, 2001).

En dépit de sa valeur nutritive, de ses effets positifs sur le traitement de la malnutrition sévère et de la réduction des risques de maladies cardiovasculaires, l'arachide peut représenter un risque de santé publique plus ou moins important en fonction des régions. En effet, la consommation d'arachide ou de dérivés peut provoquer chez certains individus de graves réactions allergiques. Ces réactions sont dues à la présence des peptides allergènes Ara h 1, 2, 3 et 6. Les cas d'allergies liées à l'arachide sont plus importants dans les pays développés que dans les pays en voie de développement car la prévalence de la sensibilité aux allergènes précités y est plus élevée (Yang, 2010). Un second facteur de risque pour la santé humaine est la contamination à l'aflatoxine de l'arachide ou de ses produits dérivés. L'aflatoxine est un métabolite d'origine fongique que l'on retrouve dans de nombreux aliments de base, notamment dans l'arachide. C'est un agent cancérogène du foie et aussi un facteur de retard de croissance chez les enfants (Gong et al., 2004). Dans les pays développés, la contamination à l'aflatoxine des arachides pose moins de problèmes car la législation et les contrôles sur les limites d'exposition sont stricts. Dans les pays en voie de développement, plus particulièrement ceux d'Afrique au sud du Sahara, le contrôle est inefficace et les conséquences de la contamination des arachides à l'aflatoxine sur la santé des populations peuvent être graves.

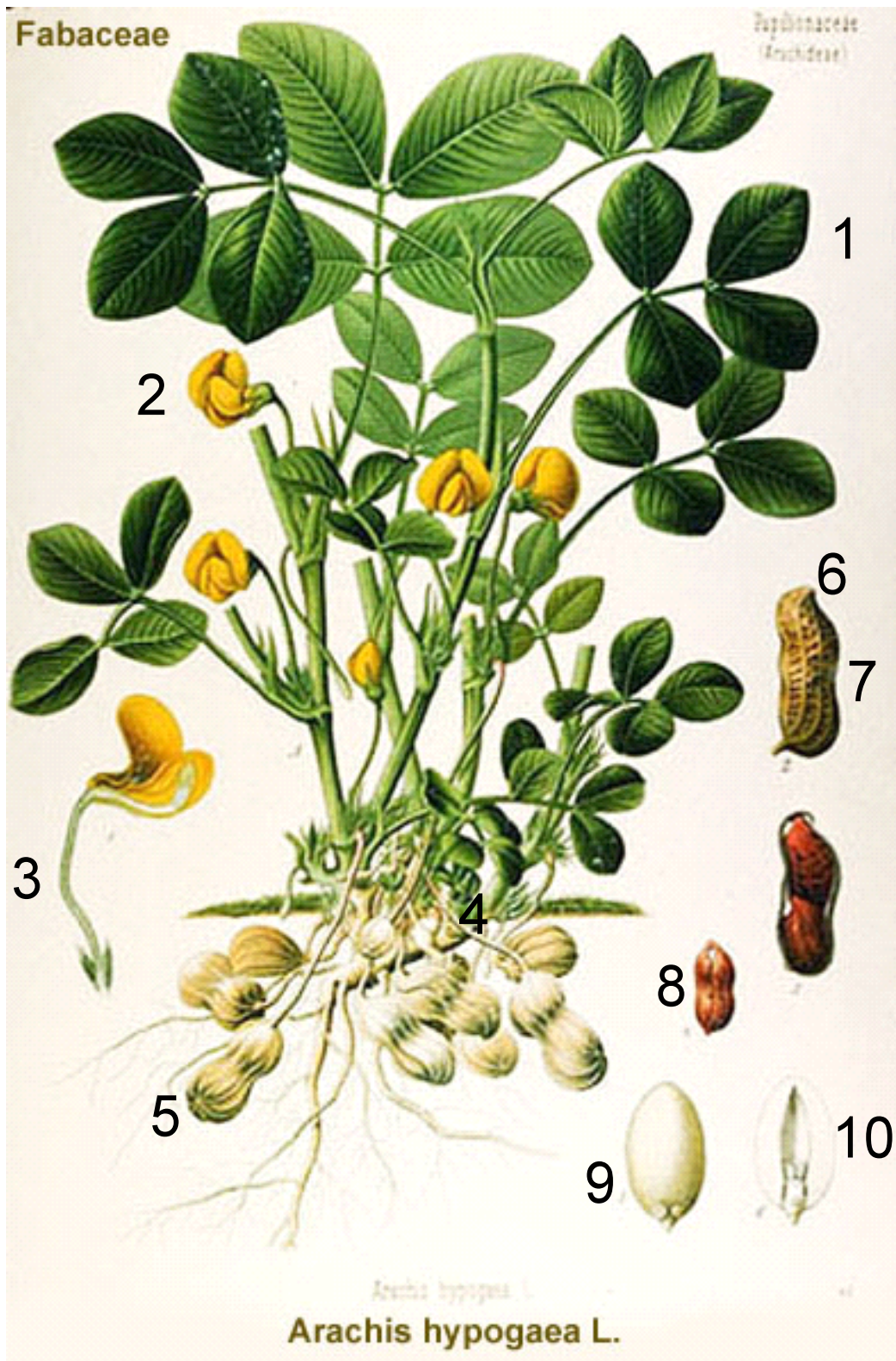


Figure I.2 : Représentation d'une plante d'arachide. 1 : feuille composée de 4 folioles, 2 : fleur, 3 : hypanthe, 4 : gynophore, 5 : gousse, 6 : bec de la gousse, 7 : constriction ; 8 : tégument de la graine, 9 : graine sans tégument, 10 : cotylédon portant l'hypocotyle, l'épicotyle et la radicule

I-2 ORIGINE ET DISTRIBUTION

L'arachide est une plante tropicale originaire d'Amérique du Sud. Le centre d'origine se situe à l'est des Andes dans une région comprise entre le sud-est de la Bolivie, le nord-ouest de l'Argentine, le nord du Paraguay et la région ouest du Matto Grosso au Brésil (Jarvis et al., 2003; Ferguson et al., 2005) (Figure I.1). La région du Matto Grosso au Brésil représente un point chaud de diversité dans lequel on retrouve la plupart des espèces sauvages du genre *Arachis*. L'espèce cultivée (*A. hypogaea* L.) a été probablement domestiquée sur le versant oriental des Andes entre le sud de la Bolivie et le nord de l'Argentine car c'est dans cette zone que l'on retrouve ses plus probables progéniteurs (Krapovickas, 1968). L'arachide cultivée a été subdivisée en deux sous-espèces *A. hypogaea ssp. hypogaea* et *A. hypogaea ssp. fastigiata*. Sept centres primaires de diversité correspondant à la répartition des deux sous-espèces et des différentes variétés botaniques de l'arachide cultivée ont été décrits. Le premier se situe dans la région du Guaraní (entre le bassin de la rivière Paraná au Paraguay et le sud-est du Brésil), le second dans les régions Goiás et Minas Gerais au Brésil, le troisième dans la région Rondônia et le nord-est du Matto Grosso au Brésil, le quatrième dans la région bolivienne (versant oriental des Andes), le cinquième dans la région péruvienne vers la côte ouest, la sixième dans le nord-est du Brésil et le septième au niveau de la région équatorienne (Gregory and Gregory, 1976a; Stalker and Simpson, 1995).

L'Afrique de l'Ouest et l'Asie du Sud-Est ont été décrites comme les deux centres secondaires de diversité de l'arachide cultivée. Le transfert de la plante en Afrique de l'Ouest s'est fait à partir des côtes brésiliennes par les navigateurs portugais et en Asie du Sud-est à partir des côtes péruviennes par les navigateurs espagnols. La diversité observée dans ces centres secondaires semble résulter des effets combinés de l'hybridation naturelle entre les différents types botaniques et de la sélection.

I-3 BIOLOGIE ET MODE DE REPRODUCTION DE L'ESPECE CULTIVEE

L'arachide est une plante annuelle, à mode de reproduction autogame. Cependant le taux d'allogamie n'est pas nul et peut varier en fonction des types botaniques et des insectes pollinisateurs. Une représentation de la plante est montrée dans la Figure I.2. Les inflorescences naissent à l'aisselle des feuilles, sur les branches primaires et secondaires. Ces inflorescences sont simples ou composées et comprennent chacune jusqu'à 5 fleurs. Une seule fleur s'ouvre par inflorescence et par jour. Les fleurs sont de type papilionacé et sont perchées à l'extrémité d'un pédoncule tubulaire allongé ou hypanthe. Les styles contenus à l'intérieur

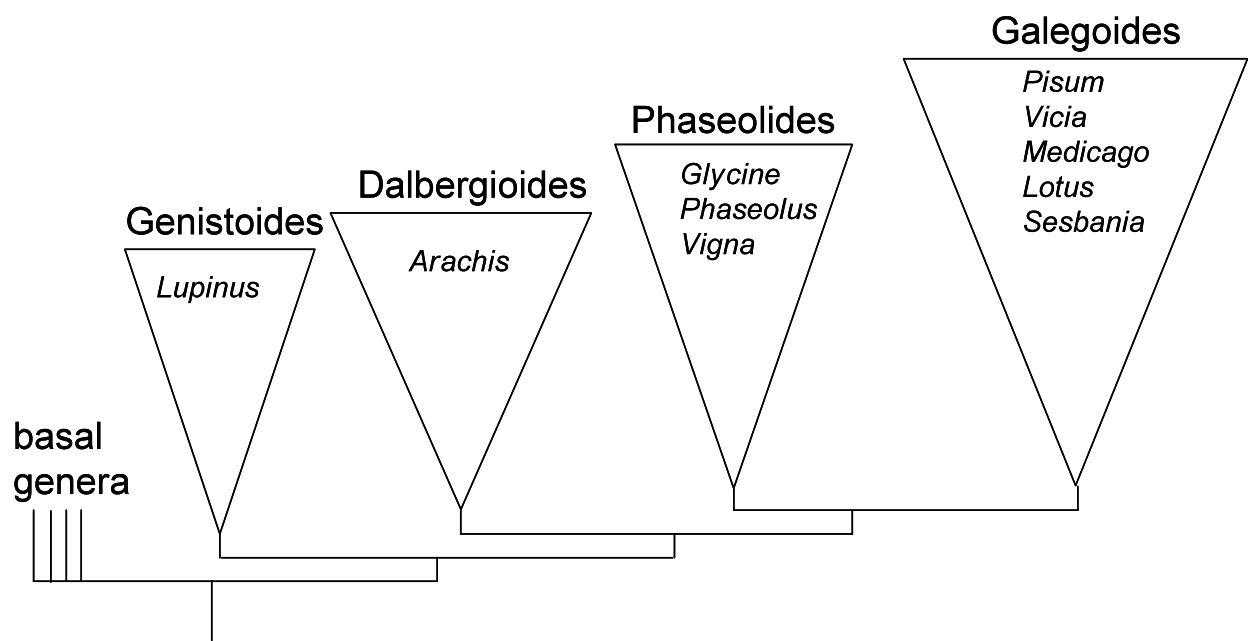


Figure I.3 : Arbre phylogénétique des Papilionoideae (Wojciechowski et al., 2003)

de l'hypanthe s'allongent en même temps que ce dernier pendant les 12 ou 24 heures précédant l'anthèse et peuvent atteindre 5 cm ou plus. La couleur des fleurs varie de l'orange foncé au jaune clair. Les fleurs contiennent 10 androcées avec 5 anthères qui s'allongent, les cinq restants devenant plus petites et globulaires. Le pollen est habituellement mature 6 à 8 heures avant l'anthèse. A maturité, le pollen est constitué de deux cellules contenant chacune deux noyaux. Les stigmates sont en général de même taille ou légèrement plus petit que les anthères. La pollinisation a lieu en même temps que l'anthèse c'est-à-dire quelques heures après le lever du soleil. Les enzymes exprimées à la surface des stigmates qui facilitent l'adhésion du pollen se dégradent dans les 8 heures qui suivent l'anthèse. C'est pour cette raison qu'il est important, dans les programmes d'hybridation artificielle, de réaliser les croisements dès les premières heures de la matinée.

L'ovaire de l'arachide est uniloculaire et possède habituellement un à trois ovules. Le sac embryonnaire possède un œuf proéminent et le noyau de l'endosperme est entouré de grains d'amidon. Après la fécondation, les grains d'amidon disparaissent et un pro-embryon se développe. Les fleurs se fanent environ 24 heures après la fécondation mais l'hypanthe reste attaché à la base l'ovaire pendant environ 5 jours. Le pro-embryon au stade 8 ou 16 cellules devient quiescent au moment où s'active un méristème situé à la base de l'ovule. Ce méristème est responsable de la mise en place d'un gynophore doté d'un géotropisme positif qui s'allonge jusqu'à pénétrer dans le sol. Une fois dans le sol, la progression du gynophore devient horizontale, puis s'arrête et les divisions cellulaires au niveau de l'embryon reprennent à nouveau. Les gousses se développent en général en l'absence de lumière 16 à 17 jours après la pollinisation.

Les graines sont composées de 2 cotylédons, d'une hypocotyle, d'un épicotyle et d'une radicule. Les cotylédons représentent environ 96% du poids de la graine.

I-4 SYSTEMATIQUE ET TAXONOMIE

L'arachide fait partie de la grande famille des légumineuses. Cette famille est divisée en trois sous-familles : les *Mimosoideae*, les *Caesalpinoideae* et les *Papilionoideae*. C'est dans cette dernière sous-famille que l'on retrouve la plupart des légumineuses cultivées à haute importance économique. Les *Papilionoideae* peuvent être subdivisés en 4 clades majeures (Figure I.3) (Wojciechowski et al., 2004) :

- Les Phaseoloides comprenant le haricot, le niébé, le soja...

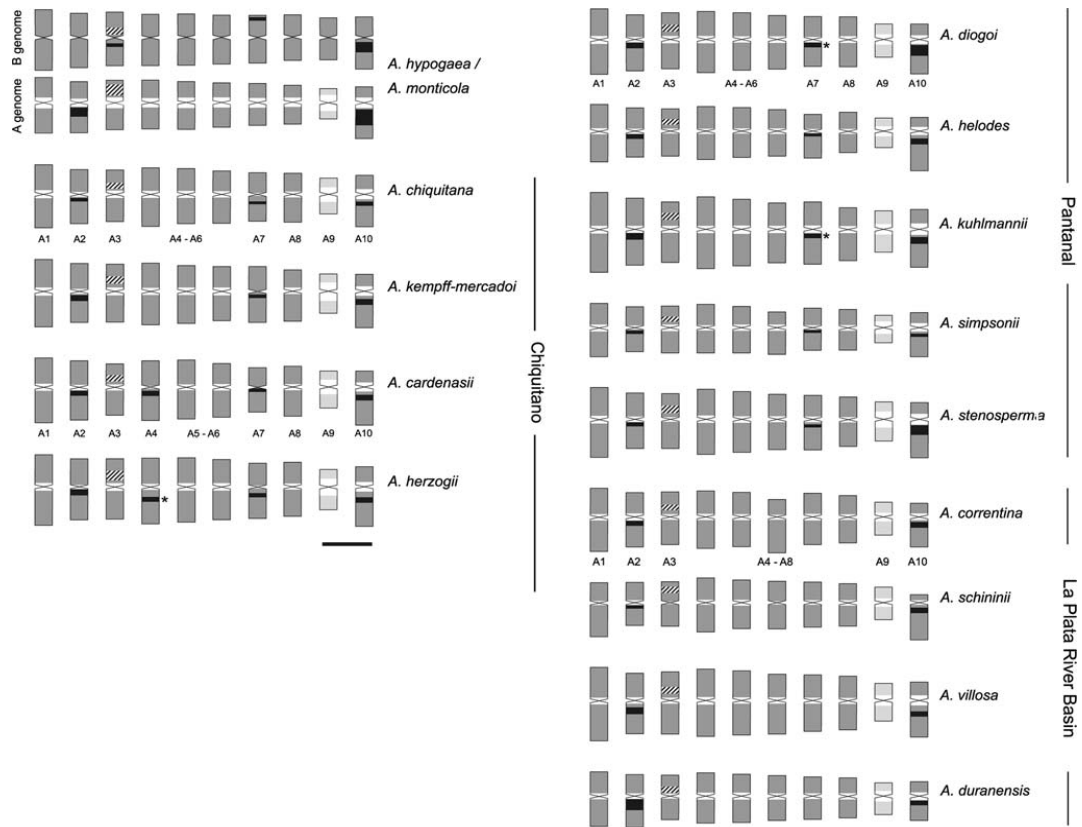


Figure I.4 : Idiogramme des espèces à génome A du genre *Arachis*. Distribution des ARN 5S (en rayé) et 18-25S (en noir) sur les chromosomes. (Robledo et al. 2009)

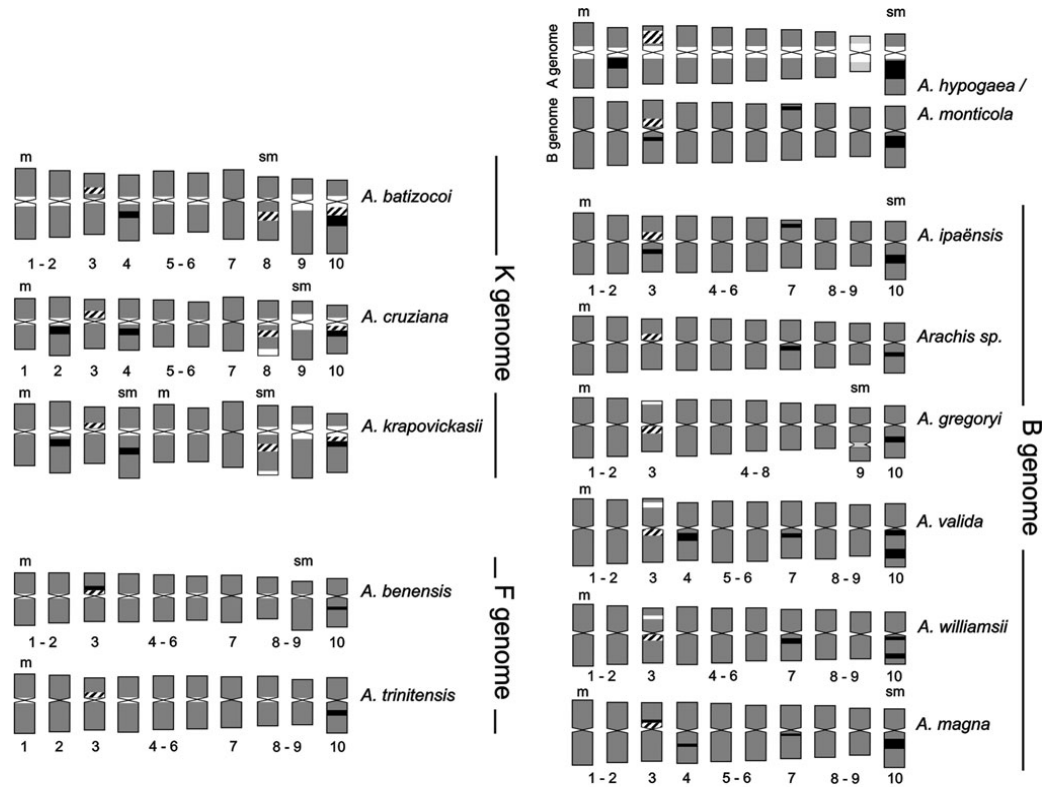


Figure I.5 : Idiogramme de la nouvelle classification des espèces à génome B du genre *Arachis*. Distribution des ARN 5S (en rayé) et 18-25S (en noir) sur les chromosomes. (Robledo et Seijo. 2010)

- Les Galegoides comprenant le pois, les lentilles, les fèves, le pois chiche et les espèces modèles *Medicago truncatula* et *Lotus japonicus*...
- Les Genistoides comprenant le lupin...
- Les Dalbergioides comprenant l'arachide...

L'arachide appartient à la tribu des *Aeschynomeneae*, la sous-tribu des *Stylosanthenae* et au genre *Arachis*. Le genre *Arachis* comprend 80 espèces décrites qui ont été réparties en 9 sections en fonction de leur morphologie, de leurs caractéristiques chromosomiques et de leur compatibilité de croisement (Krapovickas and Gregory, 1994; Valls and Simpson, 2005). Les sections *Caulorrhizae*, *Erectoides*, *Extranervosae*, *Heteranthae*, *Procumbentes*, *Trierectodes* et *Triseminatae* sont composées uniquement d'espèces diploïdes ($2n=2x=20$) (Stalker and Simpson, 1995). Les sections *Arachis* et *Rhizomatosae* sont composées d'espèces diploïdes ($2n=2x=20$, $2n=2x=18$) et d'espèces tétraploïdes ($2n=4x=40$) (Smartt and Stalker, 1982).

L'arachide cultivée appartient à la section *Arachis* dans laquelle 29 espèces diploïdes et tétraploïdes ont été décrites.

I-5 DESCRIPTION DES ESPECES DE LA SECTION ARACHIS

La section *Arachis* comprend 27 espèces sauvages diploïdes et deux espèces tétraploïdes annuelles qui sont *A. monticola* ($2n=4x=40$), la seule espèce tétraploïde sauvage de la section, et l'espèce cultivée *A. hypogaea* ($2n=4x=40$). Les espèces sauvages de cette section sont pérennes ou annuelles et ne présentent ni rhizomes ni stolons. Le système racinaire est caractérisé par une racine pivotante sans épaississement, la tige principale est érigée et les branches secondaires rampantes. Les fleurs sont de couleur orange ou jaune et les fruits sont souterrains (Krapovickas and Gregory, 1994). Le nombre chromosomique de base est $x=10$ pour toutes les espèces de cette section à l'exception des espèces diploïdes *A. palustris*, *A. decora* et *A. praecox* qui ont un nombre chromosomique de base $x=9$ ($2n=2x=18$) (Lavia, 1998; Peñaloza and Valls, 2005).

Les espèces de la section *Arachis* ont été réparties en trois différents génomes sur la base de leur caryotype (Husted, 1936; Fernández and Krapovickas, 1994). Les espèces de génome A se caractérisent par la présence d'une paire de chromosomes plus petits (Figure I.4). Cette caractéristique est absente chez les espèces de génome B qui se distinguent par ailleurs par une paire de chromosomes présentant une constriction secondaire. *A. glandulifera* se caractérise par un caryotype unique, asymétrique. C'est la seule espèce décrite de génome D (Stalker, 1991). Récemment, des études de cartographie des ARN 5S et 18-25S par la

Tableau I.1: Répartition des espèces diploïdes et tétraploïdes de la section *Arachis* en fonction du génome et du nombre de chromosome

Espèces Section Arachis	Génome	N. chromosome	Cycle
<i>cardenasii</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>correntina</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>decora</i> Krapov., W. C. Gregory & Valls	A	2n=2x=18	A
<i>diogoi</i> Hoehne	A	2n=2x=20	P
<i>duranensis</i> Krapov. & W. C. Gregory	A	2n=2x=20	A
<i>helodes</i> Martius ex Krapov. & Rigoni	A	2n=2x=20	P
<i>herzogii</i> Krapov., W. C. Gregory & C. E. Simpson	A	2n=2x=20	P
<i>hoehnei</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>kempff-mercadoi</i> Krapov., W. C. Gregory & C. E. Simpson	A	2n=2x=20	P
<i>kuhlmannii</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>microsperma</i> Krapov., W. C. Gregory & Valls	A	2n=2x=20	P
<i>palustris</i> Krapov. & W. C. Gregory	A	2n=2x=18	A
<i>praecox</i> Krapov., W. C. Gregory & Valls	A	2n=2x=18	A
<i>simpsonii</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>stenosperma</i> Krapov. & W. C. Gregory	A	2n=2x=20	P
<i>villosa</i> Benth.	A	2n=2x=20	P
<i>hypogaea</i> L.	AB	2n=4x=40	A
<i>monticola</i> Krapov. & Rigoni	AB	2n=4x=40	A
<i>batizocoi</i> Krapov. & W. C. Gregory	B, K?	2n=2x=20	A
<i>benensis</i> Krapov., W. C. Gregory & C. E. Simpson	B, F?	2n=2x=20	A
<i>cruziana</i> Krapov., W. C. Gregory & C. E. Simpson	B, K?	2n=2x=20	A
<i>gregoryi</i> C. E. Simpson, Krapov. & Valls	B	2n=2x=20	-
<i>ipaensis</i> Krapov. & W. C. Gregory	B	2n=2x=20	A
<i>krapovickasii</i> C. E. Simpson, D. E. Williams, Valls & I. G. Vargas	B, K?	2n=2x=20	-
<i>magna</i> Krapov., W. C. Gregory & C. E. Simpson	B	2n=2x=20	A
<i>trinitensis</i> Krapov. & W. C. Gregory	B, F?	2n=2x=20	A
<i>valida</i> Krapov. & W. C. Gregory	B	2n=2x=20	A
<i>williamsii</i> Krapov. & W. C. Gregory	B	2n=2x=20	A
<i>glandulifera</i> Stalker	D	2n=2x=20	A

technique FISH ont permis de mettre en évidence des différences au sein des espèces de génome B et de proposer deux génomes supplémentaires : les génomes K et F (Figure I.5) (Robledo and Seijo, 2010).

L'analyse des caryotypes des espèces tétraploïdes *A. monticola* et *A. hypogaea* a révélé 20 chromosomes à caryotype similaire à celui du génome A et 20 autres chromosomes à caryotype similaire à celui du génome B (Husted, 1936; Kochert et al., 1996; Seijo et al., 2007). Ces deux espèces ont une structure allotétraploïde AABB.

Une répartition des espèces en fonction de leur nombre chromosomique de base, de leur cycle et de leur génome est présentée en Tableau I.1.

I-6 RELATION ENTRE ESPECES DIPLOÏDES DE LA SECTION ARACHIS

Les premières études des relations entre espèces de la section ont été réalisées par analyse de la fertilité des hybrides interspécifiques. Ces études ont permis d'établir et de confirmer l'appartenance des différentes espèces à cette section. Cependant, les mécanismes d'isolement reproductif développés par un certain nombre d'espèces et d'accessions au sein d'une même espèce ont freiné les efforts de classification. En effet, Gregory et Gregory (1979) ont rapporté que certains hybrides issus de croisements intraspécifiques avaient des pourcentages de stérilité plus importants que des hybrides issus de croisements interspécifiques. L'analyse de la diversité génétique à l'aide de marqueurs moléculaires a permis de préciser la structuration décrite à partir de l'analyse des données agro-morphologiques et à partir des analyses cytologiques. Lu et Pickersgill (1993) ont analysé la variation existante entre quelques espèces diploïdes à l'aide de 11 isozymes. Cette étude a permis de différencier la seule espèce de génome B alors décrite (*A. batizocoi*) de celles de génome A et de regrouper les espèces de génome A en fonction de leur cycle (annuelle/pérenne). La séparation entre les espèces de génome A et B a aussi été révélée avec des marqueurs RFLP (Gimenes et al., 2002). Dans cette étude, une forte similarité a été mise en évidence entre les espèces de génome A (*A. cardenasii*, *A. correntina*, *A. duranensis*, *A. kempff-mercadoi*) et une faible similarité entre celles de génome B (*A. batizocoi*, *A. ipaënsis*, *A. magna*). Une étude plus large de diversité génétique à l'aide de marqueurs AFLP a été réalisée sur un échantillon représentant la quasi-totalité des espèces diploïdes et tétraploïdes de la section *Arachis* (Milla et al., 2005). Cette étude a permis de différencier les espèces à 18 chromosomes de celles à 20 chromosomes au sein du groupe d'espèces à génome A, de confirmer la différence entre les espèces de génome A et celles de génome B, et de mettre en évidence une proximité des



Figure I.6 : Diversité génétique des espèces du genre *Arachis*, section *Arachis* (Milla et al. 2005). Représentation arborée construite à l'aide du polymorphisme de marqueurs AFLP. mont, hypo : tétraploïde AB, bati, cruz, hoeh, ipae, magn, vali, will : diploïdes génome B, glan : diploïde génome D, autres : diploïdes génome A.

espèces à génomes B et D (Figure I.6). Cette étude a aussi permis de montrer une proximité génétique entre les espèces suivantes :

- *A. stenosperma*, *A. kuhlmanni*, *A. helodes*, *A. kempff-mercadoi*, *A. herzogii*, *A. diogoi*, *A. cardenasii*, *A. villosa*, *A. correntina* et *A. duranensis* (génome A, $2n=2x=20$)
- *A. ipaensis*, *A. magna*, *A. valida* et *A. williamsii* (génome B, $2n=2x=20$)
- *A. palustris*, *A. praecox* et *A. decora* (génome A, $2n=2x=18$)
- *A. benensis* et *A. trinitensis* (génome A, $2n=2x=20$)
- *A. batizocoi* et *A. cruziana* (génome B, $2n=2x=20$)

Si la classification des espèces à génome A est globalement admise, celle des espèces à génome B soulève encore quelques questions. Ceci est principalement lié à la faible similarité génétique entre espèces et, comme mentionné ci-dessus, aux différences observées sur la base de la cartographie des ARN 5S, 18-25S : ce qui a valu une proposition de subdivision des espèces en génome B, F et K. Cependant, dans une étude récente, Burow et al. (2009) ont proposé une classification des espèces de génome B en 2 groupes sur la base de l'analyse de profil RFLP : un premier groupe à profil RFLP proche de *A. batizocoi* comprenant *A. cruziana* et *A. krapovickasii*, et un second groupe à profil proche de *A. ipaensis* comprenant *A. magna* et *A. gregoryi*.

I-7 LES ESPECES DIPLOÏDES SAUVAGES A L'ORIGINE DE L'ESPECE CULTIVEE

L'identification des espèces diploïdes sauvages à l'origine de l'espèce cultivée a fait l'objet de nombreuses études durant ces 80 dernières années. La découverte de la structure allotétraploïde de l'arachide cultivée (Husted, 1936) a été le point départ des recherches sur l'identité des génomes donneurs. Les premières expériences de croisement entre *A. hypogaea* et *A. monticola* ont permis de révéler la proximité génétique entre ces deux espèces et de proposer *A. monticola* comme le parent sauvage à partir duquel *A. hypogaea* a été domestiquée (Hammons, 1970). Gregory et Gregory (1976b) ont émis l'hypothèse selon laquelle l'espèce cultivée serait issue d'une hybridation interspécifique entre une espèce pérenne telle que *A. cardenasii* et une espèce annuelle telle que *A. duranensis*. De l'étude des chromosomes de plusieurs espèces sauvages, Smartt et al. (1978) ont conclu que plusieurs espèces pourraient être de potentiels donneurs du génome A mais que le génome B proviendrait de *A. batizocoi*, la seule espèce de génome B jusqu'alors décrite. La plupart des études qui ont suivi ont porté sur l'identification des potentiels donneurs de génome A, *A. batizocoi* étant globalement admis comme donneur du génome B. Murty et Jahnavi (1986) ont

analysé la morphologie des chromosomes pachytènes des hybrides F1 interspécifiques *A. correntina* x *A. hypogaea* et en ont déduit que *A. correntina* pourrait être un donneur potentiel du génome A.

Stalker et al. (1991) ont été les premiers à remettre en question l'identité du donneur du génome B. Ces auteurs ont conclu, à partir d'une analyse de diversité génétique à l'aide de marqueurs isozymes, que la distance génétique entre *A. batizocoi* et les autres espèces de la section *Arachis* était telle que cette espèce ne pouvait pas être impliquée dans l'évolution de l'espèce cultivée. De nouvelles études portant sur la comparaison des profils de migration des protéines de stockage des graines (Bianchi-Hall et al., 1993) et sur des analyses de diversité avec des marqueurs isozymes (Lu and Pickersgill, 1993; Stalker et al., 1994) ont confirmé cette hypothèse. Kochert et al. (1991) ont tenté de reconstituer les profils de migrations de bandes RFLP des espèces tétraploïdes avec celles de plusieurs espèces diploïdes. Les résultats les plus concluants étaient ceux issues de la combinaison d'*A. ipaensis* et d'*A. duranensis*. Les analyses cytologiques de plusieurs espèces sauvages combinées à des données morphologiques et de distribution géographique ont aussi abouti à l'hypothèse selon laquelle *A. ipaensis* et *A. duranensis* seraient les plus probables progéniteurs de l'espèce cultivée (Fernández and Krapovickas, 1994). Kochert et al. (1996) ont été les premiers à combiner des données de cytologie et de marquage moléculaire pour identifier les progéniteurs de l'espèce cultivée. Ces auteurs ont aussi produit un hybride interspécifique *A. ipaensis* x *A. duranensis* et comparé son profil RFLP avec celui de *A. hypogaea*. Les résultats de cette étude ont montré que l'espèce sauvage *A. ipaensis* possédait des caractéristiques caryotypiques communes aux espèces de génome B et que le profil RFLP de l'hybride interspécifique était similaire à celui de l'espèce cultivée. Ils en ont conclu que *A. ipaensis* et *A. duranensis* étaient les plus probables progéniteurs de *A. hypogaea*. De plus, dans cette même étude, *A. duranensis* a été proposé comme potentiel parent femelle sur la base de la comparaison des profils de restrictions de l'ADN chloroplastique des trois espèces. Plus récemment, les résultats des études de cartographie physique des ARN 5S et 18-25S par technique FISH (Seijo et al., 2004), de compatibilité entre *A. hypogaea* et le tétraploïde synthétique issu du croisement *A. ipaensis* x *A. duranensis* (Favero et al., 2006) et de reconstitution du profil RFLP du tétraploïde cultivé avec des profils diploïdes (Burow et al., 2009) ont permis de confirmer les espèces diploïdes sauvages *A. ipaensis* (génome B) et *A. duranensis* (génome A) comme les plus probables progéniteurs de l'espèce cultivée.

Tableau I.2 : Classification des sous-espèces d'*Arachis hypogaea*

Sous-espèce	Variété botanique	Type commercial
<i>hypogaea</i>	<i>hypogaea</i>	Virginia
		Runner
	<i>hirsuta</i>	Peruvian
<i>fastigiata</i>	<i>fastigiata</i>	Valencia
	<i>vulgaris</i>	Spanish
	<i>peruviana</i>	
	<i>aequatoriana</i>	

I-8 DIVERSITE GENETIQUE DE L'ESPECE CULTIVEE

Il existe une forte diversité agro-morphologique au sein de l'espèce cultivée. Cette forte diversité a permis de distinguer deux sous-espèces dans le compartiment cultivé : *A. hypogaea* subsp. *hypogaea* et *A. hypogaea* subsp. *fastigiata*. Ces sous-espèces se distinguent principalement par leur port, en général rampant chez *hypogaea* et érigé chez *fastigiata*, par l'absence de fleurs sur l'axe principal chez *hypogaea* et par la différence de coloration des feuilles qui sont de couleur vert-foncé chez *hypogaea* et vert-clair chez *fastigiata* (Krapovickas and Gregory, 1994). Ces deux sous-espèces ont été elles-mêmes subdivisées en plusieurs variétés botaniques qui comprennent plusieurs types commerciaux (Tableau I.2).

En total contraste avec ce qui est observé en termes de diversité agro-morphologique, l'arachide cultivée présente une faible diversité à l'échelle moléculaire. Cette faible diversité moléculaire a été expliquée par l'origine monophylétique récente de l'espèce tétraploïde cultivée (Kochert et al., 1996) et par la domestication qui s'en est suivie (Burow et al., 2001). Au cours de ces deux dernières décennies, d'importants efforts de recherche ont été investis pour le développement de marqueurs moléculaires et l'analyse de la diversité génétique de l'arachide. Kochert et al. (1991) ont utilisé des marqueurs RFLP pour étudier la variabilité génétique existante entre cultivars appartenant aux types commerciaux virginia, runner, valencia et spanish. Leurs résultats ont montré que la variation des profils RFLP entre cultivars était très faible. Dans une étude menée par Halward et al. (1992), aucune variation de profils RAPD n'a été observée entre plusieurs cultivars et variétés traditionnelles d'arachide. De même, les marqueurs DAF et AFLP, utilisés pour géotyper différentes variétés botaniques, n'ont révélé que 0.11% et 1.7% de polymorphisme respectivement (He and Prakash, 1997). Singh et al. (1998) ont expliqué le faible polymorphisme observé dans le compartiment cultivé par l'inadéquation du matériel génétique utilisé dans les différentes études et ont préconisé l'utilisation d'une méthode d'échantillonnage plus rigoureuse et l'utilisation de marqueurs de type SSR ou AFLP. Hopkins et al. (1999) ont été l'une des premières équipes à développer des marqueurs SSR et à prouver leur effectivité pour détecter de la variation moléculaire dans le compartiment cultivé. He et Prakash (2001) ont étudié la diversité génétique entre variétés botaniques à l'aide de marqueurs AFLP et leur lien avec la seule espèce tétraploïde sauvage décrite (*A. monticola*). Leurs résultats ont montré que les variétés botaniques *aequatoriana* et *peruviana* étaient plus proches de la sous-espèce *hypogaea* que de la sous-espèce *fastigiata* et que l'espèce sauvage *A. monticola* était très proche de la sous-espèce *hypogaea*. Raina et al. (2001) ont pu distinguer, avec des données combinées de marqueurs RAPD et ISSR, les différentes variétés botaniques et cultivars

appartenant à une même variété botanique. Cependant, il ressort de ces différentes études que la diversité génétique observée chez l'arachide cultivée demeure faible.

Au début des années 2000, l'utilisation des outils de la génomique a permis une forte progression du nombre de marqueurs microsatellites utilisables chez l'arachide. Ces marqueurs ont été dérivés de banques enrichies en microsatellites (He et al., 2003; Cuc et al., 2008), de banques d'ADN génomique (Ferguson et al., 2004; Moretzsohn et al., 2004; Gimenes et al., 2007), de banques d'EST (Proite et al., 2007; Moretzsohn et al., 2009). Ils ont permis de préciser la structuration de la diversité génétique dans le compartiment cultivé et de confirmer la faible diversité génétique qui existe dans ce compartiment. En dépit du faible nombre de marqueurs polymorphes dans le compartiment cultivé, les premières cartes génétiques résultant de croisements entre variétés cultivées ont été construites (Varshney et al., 2008; Hong et al., 2010). Ces cartes génétiques, encore peu saturées, ont été utilisées pour réaliser des études de cartographie de QTL impliqués dans la tolérance à la sécheresse et offrent de belles perspectives d'utilisation de marqueurs moléculaires dans les programmes d'amélioration génétique de l'arachide.

I-9 AMELIORATION GENETIQUE DE L'ARACHIDE CULTIVEE

Les programmes de sélection pour l'amélioration génétique de l'arachide ont connu plus ou moins de réussite en utilisant les différentes méthodes décrites pour l'amélioration de plantes autogames. En effet, les sélectionneurs ont été très tôt confrontés à la faible variabilité qui existe dans le compartiment cultivé, à la difficulté de réaliser un grand nombre de croisements et au faible nombre de descendants produits à chaque génération. Des programmes d'amélioration par mutations induites et sélection de mutants intéressants se sont largement développés entre les années 1950 et 1970. Cependant le matériel produit n'a pas connu un franc succès chez les producteurs et cette technique a été progressivement délaissée (Holbrook and Stalker, 2003). La sélection massale est aussi peu utilisée par les programmes de sélection car certains auteurs ont fait état de corrélations négatives entre le rendement et la tolérance aux maladies (Knaft and Wynne, 1995). La sélection généalogique est couramment utilisée dans les programmes de sélection car elle est peu coûteuse en ressources humaines pour effectuer les croisements.

Les techniques de backcross sont aussi de plus en plus utilisées pour transférer des caractères à héritabilité simple tant dans des croisements intraspécifiques qu'interspécifiques. Cette technique a été à la base de l'introgession dans un fonds génétique cultivé de la plupart des

gènes de résistance aux maladies issus du compartiment sauvage. Deux voies ont été privilégiées :

- la voie tri-hexaploïde qui consiste à croiser l'espèce cultivée tétraploïde avec un parent sauvage diploïde. Il en résulte un hybride triploïde stérile, qui après doublement du stock chromosomique à la colchicine donne des individus hexaploïdes qui peuvent être fertiles. Le retour au niveau de ploïdie $4x$ et au fond génétique du parent récurrent se fait par backcross successifs avec le parent cultivé et sélection d'un nombre limité d'individus fertiles pentaploïdes, aneuploïdes puis tétraploïdes. Garcia et al. (2006) ont analysé à l'aide de marqueurs RAPD plusieurs populations obtenues par la voie tri-hexaploïde à partir des croisements entre l'espèce cultivée *A. hypogaea* ($2n=4x=40$) et deux espèces sauvages, *A. cardenasii* (génome A, $2n=2x=20$) et *A. batizocoi* (génome B, $2n=2x=20$). Les résultats de cette étude ont montré qu'il y avait une réduction du pourcentage de génome donneur sauvage après chaque cycle de rétrocroisement. Cependant, cette méthode longue et difficile à mettre en œuvre présente aussi des inconvénients tels que la forte stérilité et l'instabilité des hybrides, la possibilité de perte de chromosomes ou de fragments chromosomiques.
- La voie amphidiploïde qui consiste à synthétiser un tétraploïde artificiel à partir de parents sauvages diploïdes et à croiser l'hybride tétraploïde avec le parent cultivé. Le retour au fonds génétique du parent récurrent se fait par backcross successifs. Cette technique présente l'avantage de générer directement des individus ayant le même niveau de ploïdie que l'espèce cultivée. Cette technique a été utilisée avec succès pour transférer dans un fonds génétique cultivé un gène de résistance aux nématodes provenant de l'espèce sauvage *A. cardenasii* (Simpson and Starr, 2001; Simpson et al., 2003). A l'heure actuelle, des équipes de l'EMBRAPA au Brésil et de l'ICRISAT en Inde se sont investies dans la production d'amphidiploïdes sauvages qui résultent de la combinaison d'espèces de génome A et d'espèces de génome B. Cette nouvelle ressource offre des perspectives intéressantes pour l'élargissement de la base génétique de l'espèce cultivée.

La sélection récurrente, bien que très indiquée pour conserver la variabilité génétique au cours du processus de sélection, a été très peu utilisée dans les programmes de sélection arachide à cause de la difficulté à réaliser un grand nombre de croisements. Cette méthode a néanmoins été utilisée avec succès par le CIRAD et l'ISRA pour produire des cultivars présentant de bonnes caractéristiques de tolérance à la sécheresse (Clavel et al., 2007). La production

d'hybrides F_1 est une stratégie de sélection très peu envisageable chez l'arachide pour les raisons évoquées ci-dessus.

Les principales cibles de la sélection sont la résistance aux maladies, la tolérance à la sécheresse, l'amélioration de la quantité et de la qualité des huiles. Les paragraphes ci-après font un bref descriptif des avancées en termes d'amélioration de l'arachide pour les cibles de sélection précitées.

I-9-1 AMELIORATION DE LA TOLERANCE A LA SECHERESSE

L'arachide est principalement cultivée en condition pluviale dans les régions des tropiques arides ou semi-arides. Comparée à d'autres plantes annuelles tropicales, elle est décrite comme une plante qui possède un bon niveau de tolérance au déficit hydrique. Dans ses aires de culture au Sahel souvent affectées par la sécheresse, c'est la culture qui génère la plus importante quantité de matière sèche à l'hectare (Clavel et al., 2005). Cependant, un déficit hydrique pendant la floraison et surtout le remplissage des graines peut avoir un effet dépressif important sur le rendement (Rao et al., 1985; Wright et al., 1991). Le développement de cultivars adaptés à des environnements caractérisés par la faiblesse et l'irrégularité des événements pluvieux est un objectif majeur des programmes de sélection.

Les méthodologies de sélection pour l'amélioration de l'arachide pour la tolérance au déficit hydrique ont été longtemps axées sur le raccourcissement des cycles de culture pour les adapter à la longueur de la saison des pluies et sur la sélection du rendement ou de ces composantes en conditions de stress. Le raccourcissement des cycles de culture par croisement avec un parent donneur de précocité a permis d'adapter la culture de l'arachide à des régions du Sahel où la longueur de la saison des pluies dépasse rarement 90 jours. Par cette méthode, de nombreuses variétés à cycle court (75 à 90 jours) ont été développées et représentent actuellement le matériel élite utilisé dans plusieurs pays d'Afrique de l'Ouest. La sélection sur le rendement en condition de stress est une méthode qui a aussi donné des résultats et qui est encore souvent utilisée dans les programmes de sélection. Cependant cette approche a été critiquée tant pour son coût élevé en terme de ressources humaines et financières que pour le manque de reproductibilité des résultats produits à cause des fortes variations génotype x environnement observées en condition de stress hydrique (Madhava et al., 2003).

Un modèle analytique simple qui décompose le rendement en composantes physiologiques théoriquement indépendantes a été proposé pour comprendre les bases physiologiques des

facteurs qui déterminent le rendement en condition de déficit hydrique. L'équation du modèle est la suivante :

$$Y = W \times WUE \times HI$$

où Y= rendement, W= quantité d'eau évaporée par la culture, WUE= efficacité d'utilisation de l'eau par la culture (en unité de matière sèche produite par unité d'eau consommée) et HI= indice de récolte (Passioura, 1977).

Sur la base de l'équation précédente, une réécriture qui rend mieux compte des pertes d'eau liées à la transpiration de la plante que de celles liées à l'évaporation a été proposée par Fisher (1981) :

$$Y = T \times TE \times HI$$

où T est la transpiration et TE son efficacité.

Ce modèle a été largement appliquée à l'amélioration de la tolérance à la sécheresse des légumineuses à graines en général et à l'arachide en particulier (Wright and Rachaputi, 2003). Cependant, les mesures directes de transpiration ou d'extraction d'eau du sol par les racines sont assez difficiles à mettre en œuvre et surtout très peu applicables à grande échelle. La plupart des travaux ont donc porté sur la recherche de caractères de substitution qui rendent le mieux compte de l'efficacité de la transpiration chez l'arachide. Faquhar et al. (1982) ont proposé une méthode indirecte d'estimation de la transpiration par la mesure de la discrimination isotopique du carbone dans les feuilles. Chez l'arachide la discrimination isotopique du carbone est négativement corrélée à l'efficacité de la transpiration (Wright et al., 1994). Cette méthode est rapide mais assez prohibitive en termes de coût et donc difficilement applicable à un nombre important de géotypes. La mesure de la surface foliaire spécifique (SLA) a aussi été proposée comme mesure de substitution parce qu'elle est fortement corrélée à la discrimination isotopique du carbone chez l'arachide (Nageswara Rao and Wright, 1994). Cette méthode, bien qu'efficace en terme de réduction des coûts et de production de résultats, reste assujettie à de fortes variations environnementales (heure de collecte des feuilles) et physiologiques (âge des feuilles). Malgré la sensibilité de la mesure de la surface foliaire spécifique aux variations environnementales et physiologiques, ce caractère reste un bon critère de sélection pour l'amélioration des rendements en condition de stress hydrique (Nageswara Rao et al., 2000). Une bonne corrélation entre les valeurs de SPAD, de

la SLA et celles de l'efficacité d'utilisation de l'eau (WUE) a été rapportée par plusieurs auteurs (Nageswara Rao et al., 2001; Upadhyaya, 2005). De plus, une bonne corrélation entre les valeurs de SPAD, le rendement en gousses et la taille des graines sous stress a aussi été observée (Upadhyaya, 2005; Songsri et al., 2008).

Récemment, une équipe de l'ICRISAT a mis en place un système gravimétrique performant qui permet, par des pesées à différents temps correspondant à des états hydriques différents, d'avoir des estimations de l'efficacité de la transpiration d'eau par les feuilles. Cette technique est applicable à un nombre important d'individus ce qui la rend attrayante pour les programmes de sélection. Ce système a été utilisé pour étudier la variation de l'efficacité de la transpiration (TE), du SLA et du SPAD dans une population de 318 lignées recombinantes issue du croisement entre deux variétés cultivées (Krishnamurthy et al., 2007). Les auteurs de cette étude ont montré que, contrairement aux études précédentes, la corrélation entre TE, SLA et SPAD était faible. Néanmoins, Varshney et al. (2008) ont produit une carte génétique à partir de la même population et ont utilisé ces données phénotypiques dans une approche de cartographie de QTL. Ces travaux ont permis d'identifier plusieurs régions génomiques impliquées dans la variation du TE, du SPAD et du SLA et ont mis en évidence une co-localisation des QTL pour au moins deux caractères au niveau de 7 groupes de liaison.

Les indices de réponse à la sécheresse ; indice de sensibilité (SSI) (Fischer and Maurer, 1978) et indice de tolérance (STI) (Fernandez, 1993) ont été utilisés pour identifier des génotypes à rendement stable en condition de déficit hydrique. Ces indices se définissent par les équations suivantes :

$$SSI = 1 - (Y_s/Y_i) / 1 - (\tilde{Y}_s/\tilde{Y}_i)$$

$$STI = (Y_i Y_s) / \tilde{Y}_i^2$$

où Y_s et Y_i sont les rendements en gousses dans les conditions stressée et irriguée respectivement tandis que \tilde{Y}_s et \tilde{Y}_i sont les rendements moyens, en gousses, de la population dans les conditions stressée et irriguée respectivement.

L'indice de sensibilité à la sécheresse n'a pas été pertinent pour différencier les génotypes sensibles des génotypes tolérants (Clavel et al., 2005). Par contre l'indice de tolérance à la sécheresse est un paramètre déterminant pour l'identification de génotypes qui montrent une bonne performance dans des conditions de régime hydrique normales et stressées (Clavel et al., 2005; Songsri et al., 2008).

I-9-2 AMELIORATION DE LA RESISTANCE AUX MALADIES

Les maladies sont l'une des principales causes de perte de rendement dans la plupart des zones de culture de l'arachide dans le monde. Aux Etats-Unis par exemple, les pertes économiques causées par la réduction des rendements sous l'action des maladies et les investissements en actions préventives pour contrôler ces maladies se chiffrent à plusieurs millions de dollars par année (Compendium of Peanut Diseases, (1997)). Certaines pratiques culturales comme l'élimination des plantes potentiellement hôtes de certains pathogènes et la rotation des cultures, les luttés par épandage de produits chimiques et l'utilisation de génotypes résistants ont permis de contrôler voir d'éliminer certaines maladies. Cependant il existe encore un grand nombre de maladies pour lesquelles on ne trouve ni contrôle chimique efficace, ni génotypes présentant un bon niveau de résistance. La sélection de génotypes résistants aux maladies est l'un des objectifs majeurs des programmes de sélection de l'arachide dans les grands pays producteurs. Nous donnons dans les paragraphes qui suivent quelques exemples de recherche pour la sélection de cultivars résistants aux principales maladies de l'arachide.

La résistance aux maladies foliaires. A l'échelle mondiale, les cercosporioses causées par *Cercospora arachidicola* et *Cercosporidium personatum* et la rouille causée par *Puccinia arachidis* sont les maladies les plus destructrices de l'arachide (Shokes and Culbreath, 1997). Elles peuvent causer dans certaines zones jusqu'à 70% de perte de rendement. Plusieurs sources de résistance aux cercosporioses ont été identifiées dans le compartiment cultivé de l'arachide (Holbrook and Stalker, 2003). Ces sources ont été largement utilisées pour développer plusieurs cultivars. Cependant ces cultivars ne présentent qu'une résistance partielle à ces maladies. Ceci est principalement lié à la complexité du caractère de résistance aux cercosporioses qui fait intervenir plusieurs composantes potentiellement indépendantes telles que l'infection initiale, la taille des lésions, la sporulation et la défoliation (Green and Wynne, 1986; Anderson et al., 1993; Waliyar et al., 1995). Des niveaux élevés de résistance voire une immunité aux cercosporioses ont été observés chez certaines espèces diploïdes sauvages apparentées à l'arachide cultivée (Stalker and Simpson, 1995). Une étude récente a permis de cartographier plusieurs QTL de résistance à la cercosporiose tardives à partir d'un croisement en deux espèces diploïdes sauvages *A. duranensis* et *A. stenosperma* (Leal-Bertioli et al., 2009). Des programmes de recherche pour l'introgession de ces gènes dans un fonds génétique cultivé sont en cours dans plusieurs pays (Brésil, Inde, Etats-Unis...) mais sont

confrontés à la difficulté de se départir des caractères négatifs, tels que la réduction du rendement, apportés par les espèces sauvages (Holbrook and Stalker, 2003).

En ce qui concerne la résistance à la rouille, entre 1977 et 1985, un grand programme de criblage pour l'identification de génotype résistant a été réalisé sur un échantillon d'environ 10000 accessions de la collection d'arachides cultivées de l'ICRISAT (Subrahmanyam et al., 1989). Cette étude a montré que les accessions présentant un bon niveau de résistance à la rouille appartenaient majoritairement à la sous-espèce *fastigiata*. Mehan et al. (1996) ont criblé 970 nouvelles accessions et ont confirmé la meilleure résistance à la rouille des accessions appartenant à la sous-espèce *fastigiata*. Cependant, à l'instar de ce qui a été observé pour les cercosporioses, la complexité du caractère de résistance à la rouille explique le niveau partiel de résistance des accessions identifiées. Des niveaux plus élevés de résistance ont été observés chez des hybrides interspécifiques issus des croisements entre *A. hypogaea* et les espèces diploïdes sauvages de la section *Arachis* : *A. villosa*, *A. correntina*, *A. duranensis*, *A. stenosperma*, *A. batizocoi* et *A. cardenasii* (Dwivedi et al., 2002). De même, Mallikarjuna et Sastri (2002) ont observé une bonne résistance chez des hybrides issus du croisement entre *A. hypogaea* et *A. glabrata* (section *Rhizomatosae*).

Résistance à la maladie de la rosette. La maladie de la rosette est causée par un complexe composé de deux virus (le Groundnut Rosette Virus (GRV), un umbravirus et le Groundnut Rosette Assistor Virus (GRAV), un luteovirus) et d'un ARN satellite (Compendium of Peanut Diseases,(1997)). L'infection par le GRV ou le GRAV ne provoque pas de symptômes. Les formes plus ou moins sévères de la maladie sont dues aux différentes variantes de l'ARN satellite. Pour que la maladie se développe il faut que les trois éléments du complexe soient réunis dans une même plante (Naidu et al., 1999). Cette maladie sévit principalement dans les zones de culture de l'arachide en Afrique au sud Sahara et peut causer d'importantes pertes de rendement. Au Nigéria par exemple, l'épidémie de rosette de 1975 a causé des pertes estimées à environ 250 millions de dollars. Contrairement aux cercosporioses et à la rouille, de bonnes sources de résistances ont été identifiées au sein des ressources génétiques cultivées de l'arachide et sont utilisées pour développer des cultivars résistants à cette maladie (Subrahmanyam et al., 1998).

Résistance aux nématodes. Les nématodes du genre *Meloidogyne* sont ceux qui provoquent les plus fortes pertes de rendement chez l'arachide. Dans les zones de culture infestées par les nématodes, les pertes de rendement peuvent atteindre 90% (Compendium of Peanut Diseases,

1997). Les nématodes *M. hapla* et *M. arenaria* attaquent tous les organes souterrains de l'arachide (racines, gynophores et gousses). La sélection de cultivars résistants aux nématodes a fait l'objet de nombreuses recherches aux Etats-Unis. Les premières expériences de criblage ressources génétiques cultivées ont permis d'identifier plusieurs génotypes qui présentaient des degrés variables de résistance (Holbrook and Noe, 1992; Holbrook et al., 1996). Comparés aux cultivars sensibles, ces génotypes présentaient une réduction du nombre de gales, du poids et du nombre d'œufs par système racinaire. Des niveaux élevés de résistance à *M. arenaria* ont été identifiés chez les espèces sauvages apparentées à l'arachide cultivée. *A. cardenasii* ($2n=2x=20$) a été l'une des premières sources utilisées pour transférer, à l'aide d'une stratégie de backcross, la résistance aux nématodes dans un fonds génétique cultivé (Holbrook and Stalker, 2003). L'analyse de la ségrégation de la résistance dans une population F_2 entre un descendant du croisement précité et une lignée sensible a permis d'identifier deux gènes dominants responsables de la résistance. L'introgression de la résistance à *M. arenaria* a aussi été réalisée à partir du croisement entre un hybride interspécifique complexe, réunissant des sources de résistance de trois espèces sauvages [*(A. cardenasii* x *A. diogoï*) x (*A. batizocoi*)]^{x4}, et une variété cultivée. Un programme de backcross a permis l'introgression des gènes de résistance de ces sources dans des populations de sélection. Ce programme a abouti à la création de deux variétés d'arachide : COAN (Simpson and Starr, 2001) et NemaTam (Simpson et al., 2003), qui possèdent des niveaux élevés de résistance aux nématodes. L'introgression des gènes de résistance aux nématodes des espèces sauvages dans un fonds génétique cultivé est à ce jour l'une des expériences la plus aboutie dans le domaine de l'utilisation des espèces sauvages pour l'amélioration de l'arachide cultivée.

I-9-3 AMELIORATION DE LA QUALITE DES HUILES ET DE LA QUALITE GUSTATIVE

La qualité des huiles et la qualité gustative de l'arachide sont d'une importance économique capitale autant pour les huileries que pour la confiserie. Les propriétés de l'huile d'arachide sont déterminées par la composition en acide gras. L'acide oléique (O) et l'acide linoléique (L) représentent environ 80% de la composition en acides gras de l'huile d'arachide (Holbrook and Stalker, 2003). L'acide linoléique étant moins saturé et donc moins stable que l'acide oléique, une huile d'arachide de bonne qualité se caractérise par un rapport O/L élevé. La composition standard en O et L des huiles de la plupart des cultivars est de 55% et 25% respectivement (Knauff et al., 1993). Cependant, un criblage d'environ 500 génotypes sur la

composition en acides gras a permis de distinguer deux lignées possédant 80% de O et seulement 2% de L (Norden et al., 1987). La découverte d'une forte variabilité sur la teneur en acide oléique de l'huile a été le point de départ pour la sélection de cultivars ayant un ratio O/L élevé. Les travaux de Moore et Knauff (1989) ont montré que le caractère « forte teneur en acide oléique » est sous le contrôle de gènes récessifs dupliqués ol_1 et ol_2 . Lopez et al. (2001) a étudié la ségrégation du caractère dans des variétés de type spanish et a aussi conclu que deux gènes majeurs étaient responsables de la teneur élevée en O mais que d'autres gènes à effet mineur pourraient avoir un rôle dans l'expression de ce caractère. Sur la base des connaissances disponibles sur la génétique de ce caractère, plusieurs variétés commerciales à forte teneur en acide oléique ont été développées. C'est le cas par exemple de la variété commerciale SunOleic95R qui possède un ratio O/L supérieur à la plupart des variétés de type runner utilisées aux Etats-Unis (Holbrook and Stalker, 2003).

L'amélioration de la qualité gustative de l'arachide grillée est un objectif important des programmes de sélection. Les travaux de Sanders et al. (1995) ont montré que la qualité gustative de l'arachide grillée est un caractère héritable. Sur la base de ces travaux, la qualité gustative de tous les cultivars développés aux Etats-Unis depuis les années 1930 a été analysée (Isleib et al., 2000). Les résultats de ces analyses ont montré que les cultivars de type virginia étaient de moindre qualité gustative que ceux de type spanish et de type runner. L'analyse des pedigrees des cultivars de type virginia a montré qu'ils partageaient pour la plupart d'entre eux un ancêtre commun, 'Jenkins Jumbo', utilisé pour améliorer la taille des graines. La faible qualité gustative de ces cultivars pourrait provenir de cet ancêtre (Isleib et al., 1995). Pattee et al. (2001) ont rapporté que certaines lignées, sources de résistance à des maladies, créées avec des ressources génétiques d'origine exotique, peuvent réduire la qualité gustative des arachides grillées. Cet auteur a préconisé d'évaluer la qualité gustative des arachides précocement dans les programmes d'hybridation interspécifique.

II. UTILISATION DES RESSOURCES GENETIQUES SAUVAGES POUR L'AMELIORATION VARIETALE

II-1 DOMESTICATION ET REDUCTION DE LA VARIABILITE GENETIQUE DES ESPECES CULTIVEES

Il y a environ 10000 ans, certaines espèces végétales sauvages ont été soustraites de la sélection naturelle pour être soumises à la sélection humaine afin de produire en quantité et qualité les matières premières utiles à l'homme. Ce processus, connu sous le nom de domestication, est à l'origine de l'extraordinaire modification du mode de vie des communautés humaines et des écosystèmes terrestres. Les espèces végétales domestiquées ont été progressivement dispersées dans des milieux très variés, bien au delà de leurs centres de domestication. Elles ont évolué et se sont adaptées progressivement à des environnements différents. L'action exercée par les premières communautés humaines sur les espèces cultivées au cours des processus de domestication (sélection de caractères intéressants et reproduction des seuls individus possédant ces caractères) a modifié substantiellement la morphologie et la productivité des plantes. Cependant, certains allèles potentiellement intéressants ont pu être contre-sélectionnés au cours de ce processus (Kovach and McCouch, 2008). Les pertes de variabilité associées à la domestication superposées aux pratiques modernes de l'agriculture (homogénéisation des types variétaux, restriction des croisements à un pool génétique élite) ont conduit à une réduction sensible de la base génétique des espèces cultivées. De plus, la dispersion des espèces cultivées loin de leur centre d'origine et, dans certains cas, la domestication d'espèces à niveau de ploïdie différent de celui de leurs ancêtres ont abouti à une réduction voire une interruption des échanges géniques avec leurs parents sauvages avec pour conséquence une diminution significative de la dynamique évolutive qui favorise le maintien d'une large diversité.

La réduction de la variabilité génétique des plantes cultivées est un des principaux facteurs de vulnérabilité des systèmes de production actuels car elle compromet le maintien d'un progrès agricole constant par l'utilisation d'allèles nouveaux provenant de sources végétales diverses à base génétique plus large. A la fin du XXème siècle, plusieurs initiatives nationales et internationales ont concouru à collecter et à conserver les ressources génétiques pour la plupart des espèces cultivées (Plucknett et al., 1987). Une étude récente sur la situation des ressources génétiques a montré qu'il existe environ 6 millions d'accessions dans les différentes banques de gènes et que 50% étaient constituées de lignées élites ou de matériel de

sélection, 30% de variétés traditionnelles ou d'anciens cultivars et 15% d'espèces sauvages apparentées aux espèces cultivées (Hammer, 2003). Bien que la représentation des différents pools géniques au sein des banques de gènes reste inégale, ce nombre impressionnant d'accessions reflète assez bien la prise de conscience des sociétés humaines contemporaines de l'intérêt de la conservation des ressources génétiques des plantes cultivées. Cependant il contraste aussi de manière frappante avec la faible utilisation qui en résulte. Par exemple, une étude sur le pois chiche a montré que moins de 1% des accessions maintenues à l'ICRISAT et à l'ICARDA était utilisé par les programmes de sélection pour développer de nouveaux cultivars (Upadhyaya et al., 2006). De plus, dans la plupart des programmes de sélection, les ressources génétiques exotiques sont utilisées pour l'amélioration de caractères gouvernés par un seul ou un faible nombre de gènes tels que la résistance aux maladies ou aux insectes (Tanksley and McCouch, 1997). Jusque récemment, les ressources génétiques exotiques ont été peu ou pas utilisées pour l'amélioration de caractères à génétique plus complexe tels que le rendement ou la tolérance aux stress abiotiques. La sous-utilisation des ressources génétiques sauvages pour l'amélioration de caractères complexes était liée à la difficulté qu'avaient les sélectionneurs à se départir de certains allèles défavorables apportés par les espèces sauvages.

Au début des années 1990, l'augmentation du nombre de marqueurs de l'ADN et le développement de cartes de liaison génétique ont ouvert la voie à de nouvelles applications telles que la cartographie de régions génomiques associées à la variation de caractères d'intérêts agronomiques (QTL), le clonage positionnel et la sélection assistée par marqueurs (Paterson and Wing, 1993; Tanksley, 1993). Cette avancée significative dans le domaine de la génétique moléculaire a permis de lever un certain nombre de verrous sur la compréhension des bases génétiques de l'expression des caractères complexes. En effet, les analyses QTL ont révélé que plusieurs loci peuvent être impliqués dans l'expression d'un caractère mais que les effets associés à chaque locus sont variables, qu'en général une part importante de la variabilité phénotypique peut être expliquée par un nombre limité de QTL et enfin, que le phénotype d'une plante n'était qu'un modeste indicateur de son potentiel génétique (Tanksley and McCouch, 1997). Ces découvertes ont été les catalyseurs du développement d'un certain nombre de stratégies d'identification d'allèles favorables provenant ressources génétiques exotiques et de leur utilisation en sélection pour l'amélioration de caractères quantitatifs d'intérêt agronomique chez les plantes cultivées.

Tableau I.3 : Synthèse des résultats d'analyse AB-QTL pour des caractères de composante du rendement et de qualité des fruits chez plusieurs espèces cultivées (d'après Grandillo et al. 2007)

Espèces	Parent donneur	P. Genotypée	P. Phenotypée	N.caractère	N. QTLs	Référence
Tomate	<i>Solanum. pimpinellifolium</i>	BC ₂	BC ₂ F ₁ &BC ₃	31	120	Tanksley et al. 1996 Fulton et al. 2002
	<i>S. peruvianum</i>	BC ₃	BC ₄	47	269	Fulton et al. 1997, 2002
	<i>S. habrochaites</i>	BC ₂	BC ₃	34	155	Bernacchi et al. 1998a Fulton et al. 2002
	<i>S. neorickii</i>	BC ₂	BC ₃	42	251	Fulton et al. 2000, 2002
	<i>S. pennellii</i>	BC ₂	BC ₂ F ₁	25	84	Frary et al. 2004
Riz	<i>Oryza. rufipogon</i>	BC ₂	BC ₂ F ₁	12	68	Xiao et al. 1999
	<i>O. rufipogon</i>	BC ₂	BC ₂ F ₂	8	25	Moncada et al. 2001
	<i>O. rufipogon</i>	BC ₂	BC ₂ F ₁ &BC ₂ F ₂	13	76	Thomson et al. 2003
	<i>O. rufipogon</i>	BC ₂	BC ₂ F ₂	12	42	Septiningsih et al. 2003a
	<i>O. rufipogon</i>	BC ₂	BC ₂ F ₂	14	23	Septiningsih et al. 2003b
Orge	<i>Hordeum vulgare ssp. spontaneum</i>	BC ₂ F ₂	BC ₂ F _{2,5} &BC ₂ F _{2,6}	13	86	Pillen et al. 2003
	<i>H. vulgare ssp. spontaneum</i>	BC ₂ F ₂	BC ₂ F _{2,5} &BC ₂ F _{2,6}	13	108	Pillen et al. 2004
	<i>H. vulgare ssp. spontaneum</i>	BC ₃ DH	BC ₃ DH	11	25	Li et al. 2005
	<i>H. vulgare ssp. spontaneum</i>	BC ₂ DH	BC ₂ DH	9	86	von Korff et al. 2006
Haricot	wild common bean (G24404)	BC ₂ F _{3,5}	BC ₂ F _{3,5}	8	41	Blair et al. 2006
Poivron	<i>Capsicum frutescens</i>	BC ₂	BC ₂ &BC ₂ S ₁	10	58	Rao et al. 2003
Blé	Polyploid synthetic (W-7984)	BC ₂ F ₂	BC ₂ F ₃	5	40	Huang et al. 2003
	Polyploid synthetic (TA4152-4)	BC ₂ F ₃	BC ₂ F _{2,4}	11	10	Narasimhamoorthi et al. 2006

II-2 LA DETECTION DE QTL A PARTIR DE CROISEMENTS INTERSPECIFIQUES ET LES APPLICATIONS POUR LA CREATION VARIETALE

II-2-1 LA CARTOGRAPHIE DE QTL DANS DES POPULATIONS DE BACKCROSS AVANCES (AB-QTL)

Tanksley et al. (1996) ont développé une stratégie de cartographie de QTL dans des populations interspécifiques issues de backcross avancés qui permet l'identification et l'introggression simultanées des allèles favorables aux QTL apportés par les espèces sauvages dans un fonds génétique cultivé. Dans cette stratégie, le marquage moléculaire et le phénotypage sont effectués sur des individus en génération avancés de BC (BC₂, BC₃). Ces générations sont caractérisées par une faible fréquence de fragments provenant du génome donneur sauvage ce qui permet de travailler avec des populations composées d'individus proches du parent récurrent cultivé. De plus, une pression de sélection contre les allèles délétères issus du parent sauvage peut parfois être exercée au cours de la construction des populations permettant ainsi de réduire les effets d'épistasie et d'augmenter la puissance de détection des QTL. Un autre avantage de la stratégie AB-QTL réside dans le nombre restreint de générations supplémentaires de BC nécessaires pour générer des lignées quasi-isogéniques (NILs) possédant un allèle favorable d'origine exotique dans un fonds génétique homogène cultivé. Ces lignées représentent un matériel génétique fixé. Elles peuvent être multipliées et caractérisées dans des environnements différents pour confirmer les effets des QTL et être utilisées comme matériel de départ pour la création variétale.

La stratégie AB-QTL a été utilisée chez plusieurs espèces pour identifier des QTL impliqués dans la variation de caractères complexes tels que le rendement et la qualité organoleptique des fruits (Tableau I.3 d'après Grandillo et al. (2007)). Dans ce chapitre, nous présentons des exemples d'application de la stratégie AB-QTL à quelques espèces : la tomate, le riz, l'orge, le blé, le haricot et le poivron.

L'exemple de la tomate. Depuis une quinzaine d'années, les programmes de sélection de la tomate ont recours aux espèces sauvages apparentées afin d'identifier de nouveaux allèles utilisables pour l'amélioration génétique de la plante. Dans cet objectif, plusieurs études AB-QTL ont été menées à partir de croisements interspécifiques impliquant la tomate cultivée (*Solanum lycopersicum*) cv. E6203 et 5 espèces sauvages du genre *Solanum* : *S. pimpinellifolium* (Tanksley et al., 1996), *S. peruvianum* (Fulton et al., 1997), *S. habrochaites* (Bernacchiet al., 1998a; Bernacchi al., 1998b), *S. neorickii* (Fulton et al., 2000) et *S. pennellii*

(Frary et al., 2004, Fulton et al., 2004). Dans chacune de ces populations, 19 à 35 caractères relatifs au rendement et à la qualité organoleptique des fruits ont été mesurés. Des allèles favorables aux QTL en provenance des espèces sauvages ont été identifiés pour un peu moins de la moitié des caractères évalués (Grandillo and Tanksley, 2005). Ces allèles favorables en provenance des espèces sauvages ont été identifiés autant pour des caractères pour lesquels les parents sauvages présentaient un phénotype plus intéressant que celui du parent cultivé (composition en solides solubles par exemple) que pour ceux pour lesquels les parents sauvages étaient largement inférieures au parent cultivé (rendement, poids et couleur des fruits) (Grandillo and Tanksley, 2005). Le pourcentage d'allèles exotiques ayant un effet positif sur la variation des caractères mesurés variaient de 3% pour le rendement en fruit rouge à 88% pour la composition en solubles solides. La comparaison du nombre de QTL à effets positifs apporté par chaque espèce sauvage a été réalisée avec dix caractères communs aux cinq études. Cette comparaison a montré que la contribution en nombre de QTL différait en fonction des espèces avec 44% de QTL à effet positif provenant de *S. pimpinellifolium*, 41% de *S. peruvianum*, 28% de *S. neorickii*, 27% de *S. pennellii* et 15% de *S. habrochaites* (Grandillo and Tanksley, 2005). Ces analyses AB-QTL sur la tomate sont les premiers exemples d'application de la sélection assistée par marqueurs pour l'exploitation du potentiel génétique des espèces sauvages dans l'objectif d'améliorer les plantes cultivées.

L'exemple du riz. Des analyses AB-QTL ont porté sur l'identification de nouveaux allèles utilisables pour l'amélioration de caractères tels que le rendement et ses composantes et la qualité du grain. Ces études ont été réalisées à partir des croisements entre l'espèce sauvage *Oryza rufipogon* utilisée comme parent donneur et plusieurs lignées élites du riz cultivé (*O. sativa*) telles que l'hybride V20/Ce64 (Xiao et al., 1996, 1998), la variété Caiapo (Moncada et al., 2001), la variété Jefferson (Thomson et al., 2003) et la variété IR64 (Septiningsih et al., 2003a; Septiningsih et al., 2003b). Dans ces études, 8 à 14 caractères relatifs au rendement et à ses composantes ont été mesurés. Pour tous ces caractères, l'espèce sauvage *O. rufipogon* est inférieure aux variétés cultivées. A l'exception de l'étude impliquant IR64 comme parent récurrent où 33% des allèles favorables aux QTL étaient d'origine exotique, dans toutes les autres études plus de 50% des allèles favorables impliqués dans la variation du rendement et de ses composantes provenaient du parent sauvage *O. rufipogon*. De plus, Xiao et al. (1996) (1998) ont trouvé que pour plus de la moitié des QTL détectés il n'y avait pas d'association avec des caractères indésirables tels que l'augmentation de la taille des plantes ou une maturation tardive. Dans cette même étude, deux QTL, pour lesquels les effets positifs étaient

associés aux allèles de l'espèce sauvage, augmentaient le poids de grains par plante d'environ 18%. Par ailleurs, dans l'étude menée par Moncada et al. (2001) les allèles d'origine exotique ont aussi contribué positivement à l'amélioration de l'adaptation au stress hydrique. Ces résultats indiquent qu'en dépit de ses médiocres performances agronomiques, l'espèce sauvage *O. rufipon*, peut améliorer le rendement de variétés élites de riz.

L'exemple de l'orge. Cinq études QTL ont été menées à partir des croisements entre trois accessions de l'espèce sauvage *Hordeum vulgare ssp. spontaneum* et des variétés cultivées *Hordeum vulgare ssp vulgare* var. Apex, var. Harry (Pillen et al., 2003, 2004), var. Scarlett (Von Korff et al., 2005, 2006) et var. Brenda (Li et al., 2005). Dans les études de Pillen et al. (2003) (2004) une même accession de l'espèce sauvage a été croisée avec deux variétés cultivées et les populations BC₂F₂ résultantes ont été évaluées pour treize caractères quantitatifs dans six environnements. Les résultats ont montré que 34% et 48% des allèles favorables aux QTL provenaient du parent sauvage. De plus, dans un des croisements, un allèle provenant de l'espèce sauvage conférait une augmentation moyenne des rendements d'environ 7.7% dans les six environnements, attestant ainsi d'une bonne stabilité des QTL dans les différents environnements. La comparaison de la position des QTL détectés dans les deux populations a montré que 26% des QTL étaient communs mettant ainsi en évidence des effets d'épistasie en fonction du fonds génétique des espèces cultivées. L'évaluation d'une population BC₂-DH, issue du croisement entre une autre accession sauvage et la variété Scarlett, a été réalisée pour neuf caractères agronomiques dans 8 environnements. Les résultats ont montré que pour 36% des QTL détectés, les allèles favorables provenaient de l'espèce sauvage. Un nombre plus faible d'allèles favorables (20%) aux QTL impliqués dans le rendement et ses composantes a été détecté dans les analyses AB-QTL menées par Li et al. (2005).

L'exemple du blé. Les analyses AB-QTL ont porté sur le croisement entre deux variétés cultivées Prinz et Karl 92 et deux hexaploïdes synthétiques développés au CIMMYT construits à partir du croisement entre une variété de blé dur tétraploïde (AABB) (*Triticum turgidum* L.) var. Altar 84 et deux accessions différentes de l'espèce diploïde sauvage (DD) *Triticum tauschii* (Huang et al., 2003; Narasimhamoorthy et al., 2006). Dans l'étude de Huang et al. (2003), 60% des QTL impliqués dans la variation du rendement et de ses composantes avaient des effets positifs associés aux allèles de l'hexaploïde synthétique. Parmi ces QTL, ces auteurs en ont identifiés 4 qui expliquaient entre 5% et 15% d'augmentation de

rendement. Dans l'étude de Narasimhamoorthy et al. (2006), les allèles favorables d'origine exotique avaient des effets positifs sur la dureté du grain, le nombre de grains par épi et le nombre de talles pour 30% des QTL identifiés.

L'exemple du poivron. Une analyse AB-QTL a été menée par Rao et al. (2003) à partir d'un croisement entre le poivron cultivé (*Capsicum annuum*) var. Maor et une accession de l'espèce sauvage *Capsicum frutescens*. Dans cette étude, seuls 10% des QTL impliqués dans la variation des composantes du rendement avaient des effets positifs associés aux allèles de l'espèce sauvage. Les auteurs de cette étude ont expliqué ce faible pourcentage par la proximité génétique entre les deux espèces.

L'exemple du haricot. Une approche AB-QTL a été développée par Blair et al. (2006) à l'aide d'une population BC₂F_{3.5} issue du croisement entre la variété commerciale Cerinza et une accession sauvage G24404. Dans cette étude, 41 QTL ont été identifiés pour les huit caractères mesurés et 14 QTL avaient des effets positifs associés aux allèles du parent sauvages.

II-2-2 LES LIGNEES D'INTROGRESSION

II-2-2-1 Potentiel pour la détection et la validation de QTL

L'un des développements sous-jacent à l'analyse AB-QTL a été la création de lignées isogéniques (NILs) qui consiste à introgresser, par rétrocroisements successifs, un fragment chromosomique d'origine exotique portant un allèle favorable à l'expression d'un caractère donné dans le fonds génétique du parent récurrent (Eshed and Zamir, 1996; Tanksley and Nelson, 1996). Avec ce type de matériel, la comparaison des performances de la (des) lignée (s) isogénique (s) à celle du parent récurrent permet de valider les QTL détectés dans l'analyse AB-QTL. Les NIL(s), qui sont fixées, peuvent être caractérisées dans plusieurs environnements offrant ainsi l'avantage de préciser les effets des allèles exotiques dans le fonds génétique du parent récurrent.

Avec la multiplication des études impliquant des croisements interspécifiques, plusieurs types de populations portant des dénominations différentes ont été développés. On peut globalement distinguer deux types de populations :

- la première est constituée de lignées dans lesquelles on retrouve plusieurs fragments chromosomiques du parent donneur. Ces fragments peuvent être à l'état homozygote

ou hétérozygote. Ce type de lignées a été dénommé Backcross Inbred Lines (BILs) (Jeuken and Lindhout, 2004) ou Backcross Recombinant Inbred Lines (BCRIL) (Monforte and Tanksley, 2000a) ou pré-IL (Grandillo et al., 2007)

- La seconde est constituée de lignées dans lesquelles on ne retrouve qu'un seul fragment chromosomique du parent donneur. Ces fragments sont généralement maintenus à l'état homozygote. Ce type de lignées a été dénommé NIL ou QTL-NIL (Eshed and Zamir, 1996; Tanksley and Nelson, 1996; Monforte and Tanksley, 2000a; Frary et al., 2003), Introgression Lines (IL) (Eshed and Zamir, 1996; Von Korff et al., 2004; Grandillo et al., 2007), Chromosome Segment Substitution Line (CSSL) (Wan et al., 2004) et Stepped Aligned Inbred Recombinant Strains (STAIRS) (Koumproglou et al., 2002). Les NILs diffèrent des CSSLs et des ILs par la taille du fragment donneur introgressé qui est en général plus réduite chez les NILs.

Les lignées d'introgression (IL) offrent l'opportunité de faire des analyses QTL au niveau d'une région spécifique du génome du parent donneur. Elles peuvent aussi permettre des analyses QTL au niveau du génome entier lorsque la totalité du génome donneur est représenté par fragments contigus dans un ensemble de lignées constituant une collection d'ILs (Zamir, 2001). Les collections de lignées d'introgression présentent un certain nombre d'avantages pour la cartographie de QTL impliqués dans la variation de caractères d'intérêt agronomique. Ces avantages sont liés au fait que 1) chaque lignée ne possédant qu'un seul et unique fragment du génome donneur est phénotypiquement très proche du parent récurrent facilitant les évaluations pour des caractères tels que le rendement, 2) il y a une optimisation de la puissance de détection de QTL à effets faibles car les variations phénotypiques entre le parent récurrent et une IL donnée sont liées au seul fragment issu du donneur, 3) les effets d'épistasie liés à d'autres fragments du génome donneur sont éliminés, 4) à l'instar des NILs, les ILs sont un matériel fixé qui peuvent faire l'objet d'une caractérisation avec plusieurs répétitions et dans plusieurs environnements ce qui permet de réduire les interactions génotype x environnement (GxE) et de préciser les effets des QTL (Grandillo et al., 2007). Le niveau de résolution qu'offre une collection d'ILs pour la cartographie de QTL est fonction de la taille du fragment introgressé dans chaque lignée et du chevauchement entre les fragments adjacents contenus dans les différentes lignées constitutives de la collection (Liu et al., 2003). En général, la taille des fragments introgressés dans les ILs (20 à 30 cm en moyenne) offre une résolution relativement faible pour la cartographie de QTL. Cependant les ILs constituent le matériel de départ à partir duquel peuvent être dérivées des NILs à taille de fragment plus petite pour une cartographie fine des QTL. La cartographie fine de QTL à partir de NILs

Tableau I.4 : Collection de lignées d'introgression construites à partir de croisements interspécifiques (d'après Grandillo et al. 2007)

Plantes	Parent donneur	N. de lignées développées	% de couverture du génome	Caractères mesurés
Tomate	<i>S. pennellii</i>	50	100	Composantes du rendement, morphologie foliaire
	<i>S. pennellii</i>	75	100	Couleur et taille des fruits, composition en métabolites
	<i>S. habrochaites</i>	99	85	-
	<i>S. lycopersicoides</i>	90	96	-
Riz	<i>O. glumaepatula</i>	59	100	-
	<i>O. glaberrima</i>	39	100	-
	<i>O. meridionalis</i>	61	100	-
	<i>O. rufipogon</i>	159	67.5	Composantes du rendement
Orge	<i>Hordeum vulgare ssp spontaneum</i>	49	98	Date d'épiaison
	<i>Hordeum vulgare ssp spontaneum</i>	43	93	Date d'épiaison
Blé	<i>synthétique 6x</i>	84	100	Composantes du rendement

permet non seulement de mieux décomposer les effets phénotypiques associés au fragment (un seul QTL ou plusieurs QTL très proches impliqués dans la variation phénotypique du même caractère) mais aussi de réduire la possibilité qu'il y ait des effets indésirables associés au fragment (Thomson et al., 2003; Frary et al., 2004). De plus, la réduction de la taille du fragment donneur chez les NILs, permet d'identifier les marqueurs les plus fortement liés au QTL et de les utiliser dans les programmes de sélection assistée par marqueurs. Enfin, les NILs sont aussi un outil puissant pour le clonage positionnel des gènes sous-jacents aux QTL (Frary et al., 2000; Fridman et al., 2004).

A la suite des analyses AB-QTL, des collections d'ILs ont été développées chez quelques espèces (Tableau I.4, d'après Grandillo et al. (2007)).

Chez la tomate par exemple, une collection de 99 ILs qui offrait une couverture d'environ 85% du génome sauvage a été développée à partir de l'espèce sauvage *S. hirsutum* (Monforte and Tanksley, 2000b). Dans cette collection d'ILs une importante variation a été observée sur des caractères tels que le rendement, la forme des feuilles, la forme, la taille et la couleur des fruits. Plusieurs QTL à effets positifs associés aux allèles de l'espèce sauvage ont été détectés (Monforte and Tanksley, 2000a; Monforte et al., 2001; Yates et al., 2004). Une collection d'ILs a aussi été développée en utilisant l'espèce sauvage *S. lycopersicoides* comme parent donneur (Chetelat and Meglic, 2000; Canady et al., 2005). Cette collection de 56 ILs offrait une couverture de 96% du génome de l'espèce sauvage.

Chez l'orge, deux collections d'ILs ont été développées à partir de la même espèce sauvage, *Hordeum vulgare ssp spontaneum*, comme parent donneur (Von Korff et al., 2004). La première collection de 49 ILs dénommée S42, avec comme fonds génétique la variété cultivée Scarlett, couvrait approximativement 98% du génome sauvage avec en moyenne 3 fragments par lignée. La seconde collection de 43 ILs dénommée T42, avec comme fonds génétique la variété Thuringia, couvrait approximativement 93% du génome sauvage avec en moyenne 2.5 fragment par lignée. La collection S42 a été récemment utilisée pour valider des QTL impliqués dans la résistance au mildiou et à la rouille (Schmalenbach et al., 2008) et pour détecter des QTL impliqués dans la variation de caractères agronomiques (Schmalenbach et al., 2009). Dans cette dernière étude, 47 QTL impliqués dans la variation du nombre d'épis au mètre carré, du nombre de grains par épi, de la date d'épiaison, la hauteur de la plante, l'indice de récolte, le poids de mille grains et le rendement ont été détectés. Plus de 40% des allèles qui conféraient une amélioration significative des caractères précités provenaient de l'espèce sauvage.

Chez le riz, un travail de construction d'une collection de 159 ILs a été réalisé à partir du croisement entre la variété cultivée Guichao 2 (*indica*) et l'espèce sauvage *Oryza rufipogon* (Tian et al., 2006). La collection d'ILs couvrait environ 67.5% du génome sauvage avec en moyenne 2 fragments homozygotes et 1 fragment hétérozygote issus du parent donneur, et une majorité de fragments de petite taille (10 cM). Les 159 ILs ont été évaluées pour sept caractères relatifs au rendement dans 2 environnements. Sur les 17 QTL détectés dans les 2 environnements, 6 QTL impliqués dans l'amélioration des caractères tels que le nombre de panicules par plante, le nombre de grains par panicule et le nombre de grains matures par plante, avaient des effets positifs associés au parent donneur sauvage *O. rufipogon*.

Chez le blé (*Triticum aestivum*), une collection de 84 ILs spécifiques du génome D a été construite à partir du croisement entre le cultivar Chinese Spring et un hexaploïde synthétique comprenant le génome D de l'espèce sauvage *Aegilops tauschii* (Pestsova et al., 2006). La totalité du génome de l'espèce sauvage *Aegilops tauschii* était représentée dans la collection d'ILs, à l'exception des régions télomériques des chromosomes 1DL, 4DL et 7DS et d'une région d'une taille inférieure à 24 cM sur le bras du chromosome 3DL. Un sous-ensemble de 52 ILs a été phénotypé pour 6 caractères agronomiques et parmi les 17 QTL détectés, 9 QTL avaient des effets positifs associés aux allèles du parent donneur sauvage. Une collection de 97 ILs a aussi été développée à partir du croisement entre le parent exotique Am3 et la variété cultivée Laizhou953 (Liu et al., 2006). Le parent exotique Am3 est un hexaploïde synthétique qui a été obtenu à partir du croisement entre *Triticum carthlicum* (AABB) et *Aegilops tauschii* (DD). La collection d'ILs couvraient environ 37.7% du génome donneur et les fragments d'origine sauvage étaient retrouvés en majorité sur les chromosomes 2D, 3B, 6B et 1D. Dans cette étude, aucun fragment donneur n'a été détecté sur le chromosome 4D. Les 97 ILs ont été évalués pour 9 caractères agronomiques et certaines lignées étaient supérieures à la variété cultivée attestant ainsi des effets positifs de l'introgession d'allèles issus du donneur sauvage. Chez la laitue, une collection de 28 ILs a été développée à partir du croisement entre la laitue cultivée (*Lactuca saliva*) cv. Olof et l'espèce sauvage *Lactuca saligna* (Jeuken and Lindhout, 2004). La collection d'ILs couvraient 96% du génome sauvage, avec 20 lignées portant un seul fragment donneur à l'état homozygote, 4 lignées portant 2 fragments à l'état homozygote et 4 autres lignées portant 1 seul fragment à l'état hétérozygote. Ces lignées ont été utilisées pour cartographier 12 caractères morphologiques.

Koumproglou et al. (2002) ont développé une ressource puissante chez *Arabidopsis thaliana* pour faciliter la détection et la cartographie fine de QTL dans des régions particulières du génome. Cette ressource consiste, à partir d'un croisement entre une lignée donneuse et une

lignée receveuse, à développer un ensemble de lignées comprenant un jeu de cinq lignées d'addition monosomique (une lignée pour chacun des chromosomes d'*Arabidopsis*) et une large collection d'ILs couvrant des régions de plus en plus fines (pouvant être inférieure à 0.5 cM) du chromosome donneur contenu dans chaque lignée d'addition. Cette ressource dénommée « Stepped Aligned Inbred Recombinant Strains » (STAIRS) permet dans un premier temps de localiser un QTL sur un chromosome donné, puis sur des régions plus petites (5 à 10 cM) et enfin sur des régions de taille inférieure à 1 cM. A chaque étape un nombre restreint de lignées est utilisé pour le phénotypage ce qui permet d'optimiser les dispositifs expérimentaux en terme de répétitions avec pour conséquence une meilleure estimation de l'effet du QTL. Les STAIRS constituent un outil puissant pour la cartographie de QTL, pour l'identification de gènes candidats et pour les études d'expression.

II-2-2-2 Potentiel pour la création variétale assistée par marqueur

Les analyses QTL effectuées chez plusieurs espèces ont montré qu'il était peu réaliste d'imaginer obtenir une amélioration substantielle des caractères à génétique complexe, tels que le rendement par exemple, à partir de l'introgession d'un seul QTL dans un fonds génétique cultivé. Le pyramidage assisté par marqueurs de plusieurs QTL, provenant d'un même parent donneur ou de plusieurs donneurs différents, dans un même fonds génétique semble être une stratégie qui offre de meilleures perspectives pour l'amélioration des plantes cultivées.

Les premières expériences de pyramidage de QTL provenant d'espèces sauvages dans un fonds génétique cultivé ont été dérivées des projets d'analyse AB-QTL chez la tomate. Ces expériences ont montré qu'une lignée résultant du pyramidage de 4 QTL issus du compartiment sauvage était supérieure pour les caractères relatifs au Brix et au rendement à plusieurs lignées commerciales et au cultivar E6203 (Grandillo et al., 2007). De même Gur et Zamir (2004) ont développé la lignée IL789 en pyramidant, par une stratégie de sélection assistée par marqueurs, trois QTL associés au rendement provenant de l'espèce sauvage *L. pennellii* dans le fonds génétique du parent récurrent M82. La lignée IL789 a ensuite été croisée avec 4 lignées testeurs. Les rendements en Brix et en fruits des hybrides résultant de ces croisements étaient de 50% supérieurs à ceux de la variété de tomate hybride commerciale BOS3155 autant en condition irriguée qu'en condition d'apport limité en eau. Ces résultats ont permis de valider les effets des QTL d'origine sauvage dans différents fonds génétiques

attestant ainsi du potentiel des allèles exotiques utilisés pour l'amélioration de la stabilité des rendements dans différents environnements.

Le pyramidage assisté par marqueurs des QTL impliqués dans l'augmentation de la production de grains (Gn1) et la réduction taille de la plante (Ph1) a aussi été réalisée à partir de lignées isogéniques dérivées d'un croisement intraspécifique chez le riz (Ashikari and Matsuoka, 2006). Les QTL Gn1 et Ph1 ont été détectés dans le croisement entre Koshihikari (*japonica*) et Habataki (*indica*). Les NILs correspondant à chaque QTL ont été croisées et des lignées qui combinaient les deux QTL dans le fonds génétique de Koshihikari ont été sélectionnées. Ces lignées, comparées au parent récurrent, montraient une augmentation de la production de grains de 23% et une réduction de la taille de la plante de 20%.

**Chapitre II GENETIC MAPPING OF WILD INTROGRESSIONS INTO
CULTIVATED PEANUT: A WAY TOWARD ENLARGING THE
GENETIC BASIS OF A RECENT ALLOTETRAPLOID**

Research article

Open Access

Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid

Daniel Foncéka¹, Tossim Hodo-Abalo^{2,3}, Ronan Rivallan¹, Issa Faye², Mbaye Ndoeye Sall³, Ousmane Ndoeye², Alessandra P Fávero⁴, David J Bertoli^{5,6}, Jean-Christophe Glaszmann¹, Brigitte Courtois¹ and Jean-Francois Rami*¹

Address: ¹Centre de coopération internationale en recherche agronomique pour le développement (Cirad), UMR Développement et Amélioration des plantes, TA A96/3, Avenue Agropolis, Montpellier, France, ²ISRA: Institut Sénégalais de Recherches Agricoles, Centre National de Recherche Agronomique, BP 53, Bambey, Sénégal, ³ISRA-CERAAS: Institut Sénégalais de Recherches Agricoles, Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse, Route de Khombole, BP 3320, Thiès, Sénégal, ⁴Embrapa Recursos Genéticos e Biotecnologia, C.P. 02372, CEP 70.770-900 Brasília, DF, Brazil, ⁵Universidade Católica de Brasília, Campus II, SGAN 916, CEP 70.790-160 Brasília, DF, Brazil and ⁶Universidade de Brasília, Campus Universitário, CEP 70.910-900 Brasília, DF, Brazil

Email: Daniel Foncéka - daniel.fonceka@cirad.fr; Tossim Hodo-Abalo - aristossim@yahoo.fr; Ronan Rivallan - ronan.rivallan@cirad.fr; Issa Faye - issafaye2001@yahoo.fr; Mbaye Ndoeye Sall - mbayesall@yahoo.fr; Ousmane Ndoeye - ousndoye@refer.sn; Alessandra P Fávero - favero@cenargen.embrapa.br; David J Bertoli - davidbertoli@unb.br; Jean-Christophe Glaszmann - jean-christophe.glaszmann@cirad.fr; Brigitte Courtois - brigitte.courtois@cirad.fr; Jean-Francois Rami* - jean-francois.rami@cirad.fr

* Corresponding author

Published: 3 August 2009

Received: 20 February 2009

BMC Plant Biology 2009, 9:103 doi:10.1186/1471-2229-9-103

Accepted: 3 August 2009

This article is available from: <http://www.biomedcentral.com/1471-2229/9/103>

© 2009 Foncéka et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Peanut (*Arachis hypogaea* L.) is widely used as a food and cash crop around the world. It is considered to be an allotetraploid ($2n = 4x = 40$) originated from a single hybridization event between two wild diploids. The most probable hypothesis gave *A. duranensis* as the wild donor of the A genome and *A. ipaënsis* as the wild donor of the B genome. A low level of molecular polymorphism is found in cultivated germplasm and up to date few genetic linkage maps have been published. The utilization of wild germplasm in breeding programs has received little attention due to the reproductive barriers between wild and cultivated species and to the technical difficulties encountered in making large number of crosses. We report here the development of a SSR based genetic map and the analysis of genome-wide segment introgressions into the background of a cultivated variety through the utilization of a synthetic amphidiploid between *A. duranensis* and *A. ipaënsis*.

Results: Two hundred ninety eight (298) loci were mapped in 21 linkage groups (LGs), spanning a total map distance of 1843.7 cM with an average distance of 6.1 cM between adjacent markers. The level of polymorphism observed between the parent of the amphidiploid and the cultivated variety is consistent with *A. duranensis* and *A. ipaënsis* being the most probable donor of the A and B genomes respectively. The synteny analysis between the A and B genomes revealed an overall good collinearity of the homeologous LGs. The comparison with the diploid and tetraploid maps shed new light on the evolutionary forces that contributed to the divergence of the A and B genome species and raised the question of the classification of the B genome species. Structural

modifications such as chromosomal segment inversions and a major translocation event prior to the tetraploidisation of the cultivated species were revealed. Marker assisted selection of BC_1F_1 and then BC_2F_1 lines carrying the desirable donor segment with the best possible return to the background of the cultivated variety provided a set of lines offering an optimal distribution of the wild introgressions.

Conclusion: The genetic map developed, allowed the synteny analysis of the A and B genomes, the comparison with diploid and tetraploid maps and the analysis of the introgression segments from the wild synthetic into the background of a cultivated variety. The material we have produced in this study should facilitate the development of advanced backcross and CSSL breeding populations for the improvement of cultivated peanut.

Background

Peanut (*Arachis hypogaea* L.) is widely used as a food and cash crop around the world. It is mainly grown by resource-poor farmers in Africa and Asia to produce edible oil, and for human and animal consumption. Peanut is a member of the *Fabaceae*, tribe *Aeschynomeneae*, subtribe *Stylosanthinae*, genus *Arachis*. In this genus, 69 diploid and tetraploid species have been described [1]. *A. hypogaea* is the only species that has been truly domesticated although several species have been cultivated for their seed or forage [2]. Cultivated peanut is considered to be an allotetraploid ($2n = 4x = 40$) originated from a single hybridization event between two wild diploids with A and B genome [3]. Several studies aimed at identifying the wild diploid ancestors of *A. hypogaea*. The wild species *A. duranensis* and *A. ipaënsis* appeared to be the best candidates for the A and B genome donors, respectively [4-6].

Polyploidy is a widespread process that played a major role in higher plants' speciation and adaptation. The stages of polyploid formation usually include reproductive isolation from the progenitors [7,8]. As for many polyploid species, cultivated peanut has experienced a genetic bottleneck which, superimposed with the effects of the domestication, has greatly narrowed the genetic diversity. The low level of DNA polymorphism between cultivated genotypes has been described by many authors [9-12]. More recently, a rate of polymorphism of 12.6% has been reported between two cultivated varieties, used as parents of a RIL population, surveyed with 1145 SSR markers [13]. The low level of polymorphism within cultivated peanut has greatly hampered the application of molecular breeding approaches for the genetic improvement of cultivated peanut. Up to date, few genetic linkage maps have been published in *Arachis*. At the diploid level, three genetic maps involving species with A and B genomes, one based on RFLP markers [14] and the other ones on SSR markers [15,16], have been produced. The A genome SSR based map has been recently extended using legume anchor markers and aligned with *Medicago* and *Lotus* genomic sequences [17]. At the tetraploid level, two genetic maps were also reported. Varshney et al. [13]

reported the detection of drought tolerance QTLs based on a cultivated \times cultivated SSR genetic map. Although the genetic map remained unsaturated, due to the low level of polymorphism between cultivated peanut varieties, QTLs have been detected attesting of the interest of molecular breeding tools in genetic improvement of peanut. Burrow et al. [18] reported the construction of a RFLP map, based on a BC_1 population deriving from a cross between a wild synthetic amphidiploid (TxAG6) and a cultivated peanut variety (Florunner). The synthetic amphidiploid, used to overcome the reproductive barriers between the wild diploids and the cultivated species, allowed the genome-wide analysis of the transmission of chromatin between wild and cultivated species of the genus *Arachis*. However, the wild parents used to create the amphidiploid (*A. batizocoi*, *A. cardenasii* and *A. diogonii*) are unlikely to be the ancestors of *A. hypogaea* [12,19-21]. The genetic mapping of populations derived from the cross between the most probable wild progenitors of *A. hypogaea* and a cultivated peanut variety has, to our knowledge, never been reported.

Genome-wide introgression of a small fraction of the wild genome species while keeping the genetic background of the cultivated is a good mean to explore the largely untapped reservoir of useful alleles of interest that remain in the wild species. This is especially interesting for species with narrow genetic basis. This approach has been widely utilized for the introgression of favourable QTL(s) for various traits in tomato [22-26], in rice [27-32], in wheat [33] and in barley [34,35]. In peanut, the reproductive barriers between wild and cultivated species, the technical difficulties encountered in making large number of crosses as well as the short period between sowing and flowering have impeded the efforts to apply a Marker Assisted Backcross (MABC) approach for the development of interspecific introgression line populations.

In this study, we report for the first time the development and the analysis of the genome-wide segment introgressions of the most probable wild progenitors of the cultivated peanut species (*A. duranensis* and *A. ipaënsis*) into the background of the cultivated Fleur 11 variety through

the construction of a SSR genetic map as well as the evaluation of the coverage and the length of the wild genome segments in a BC_1F_1 and BC_2F_1 populations. This work benefits from the recently developed synthetic amphidiploid (*A. ipaënsis* × *A. duranensis*)^{4X} [5] that made possible the interspecific introgressions.

Methods

Plant material

A panel comprising 2 wild diploid accessions (*A. duranensis* V14167 diploid AA and *A. ipaënsis* KG30076 diploid BB), a tetraploid AABB amphidiploid (*A. ipaënsis* × *A. duranensis*)^{4X}, hereafter called AiAd and a cultivated tetraploid AABB variety (Fleur 11), was used in this study. The amphidiploid was developed by Favero et al. [5] by crossing *A. ipaënsis* KG30076 (B genome) with *A. duranensis* V14167 (A genome). The resulting F_1 was doubled with colchicine to produce a fertile fixed synthetic amphidiploid. Fleur 11, a local peanut variety grown in Senegal, is a Spanish type short cycle variety, high yielding and tolerant to drought. A BC_1F_1 and a BC_2F_1 populations deriving from the cross between Fleur 11 used as female recurrent parent and the amphidiploid AiAd were produced. The BC_1F_1 and BC_2F_1 populations were developed under greenhouse conditions in Senegal in 2006 and 2008

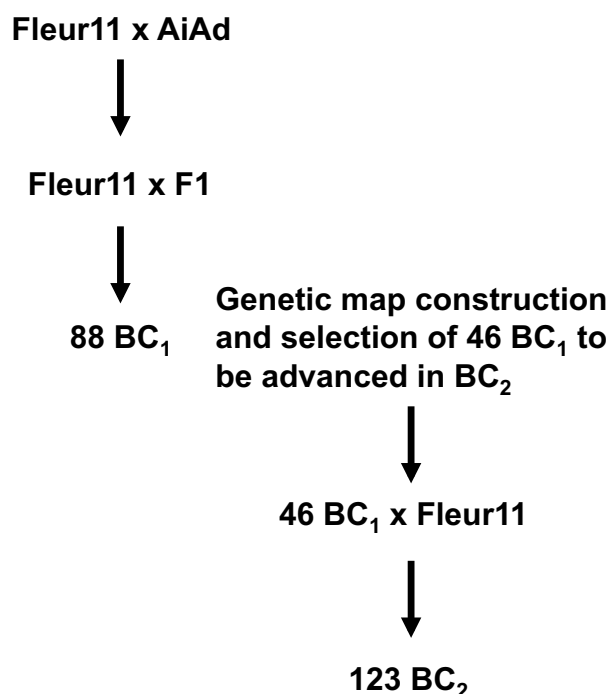


Figure 1
Breeding scheme used in the study. The cultivated Fleur 11 variety was used as female parent to produce the F_1 and the BC_1F_1 individuals, and as male parent for producing the BC_2F_1 individuals.

respectively. The crossing scheme used to generate the two populations is shown in Figure 1. The BC_1F_1 population comprised 88 individuals. Forty six BC_1F_1 plants were selected based on introgression analysis and crossed with the Fleur 11 recurrent parent to produce the BC_2F_1 generation.

DNA Isolation

Young leaves were harvested from 15 day old plants and immediately stored at 4°C in ice before DNA extraction. DNA was extracted from 100 mg of fresh leaves following a slightly modified MATAB protocol [36]. Briefly, leaves were ground in liquid nitrogen using a mortar and pestle and dissolved in 750 µL of MATAB buffer at 74°C. The samples were incubated 20 minutes at 74°C and cooled during 5 minutes at room temperature. A volume of 750 µL of CIA (24:1) was added in each sample and all samples were shaken gently until homogenization before centrifugation at 12000 rpm during 20 minutes. The supernatant was harvested and the DNA was precipitated with 600 µL of 2-propanol. After centrifugation, pellets were washed with 300 µL of 70% ethanol, air dried and dissolved in 500 µL of TE.

Microsatellite Analysis

Four hundred twenty three already-published SSR markers [12,15,21,37-45] plus 135 unpublished long size SSR markers from EMBRAPA and the Universidade Católica de Brasília were used in this study. A total of 558 SSR markers have been screened for polymorphism on the amphidiploid and its two wild diploid parents, and on the cultivated Fleur 11 variety. For a given SSR locus, the forward primer was designed with a 5'-end M13 tail (5'-CAC-GACGTTGTAAAACGAC-3'). PCR amplifications were performed in a MJ Research PTC-100™ thermocycler (Waltham, MA, USA) or in an Eppendorf Mastercycler on 25 ng of DNA in a 10 µL final volume of buffer (10 mM Tris-HCl (pH 8), 100 mM KCl, 0.05% w/v gelatin, and 2.0 mM MgCl₂) containing 0.1 µM of the M13-tailed primer, 0.1 µM of the other primer, 160 µM of dNTP, 1 U of Taq DNA polymerase (Life Technologies, USA.) and 0.1 µM of M13 primer-fluorescent dye IR700 or IR800 (MWG, Germany). The touchdown PCR programme used was as follows: initial denaturation at 95°C for 1 min; following by 10 cycles of 94°C for 30 s, T_m (+5°C, -0.5°C/cycle) for 1 min, and 72°C for 1 min. After these cycles, an additional round of 25 cycles of 94°C for 30 s, T_m for 1 min, and 72°C for 1 min and a final elongation step at 72°C for 8 min was performed. IR700 or IR800-labeled PCR products were diluted 7-fold and 5-fold respectively, subjected to electrophoresis in a 6.5% polyacrylamide gel and then sized by the IR fluorescence scanning system of the sequencer (LI-COR, USA). Migration images were analysed using Jelly 0.1 (Rami, unpublished) and exported as a data table. Segregations were checked for distortion to

the expected 1:1 ratio using a Chi² test at a significance level of 0.05.

Genetic map construction

The polymorphic markers were used to genotype 88 individuals of the BC₁F₁ population. The linkage analysis was performed using Mapdisto software version 1.7.2.4 [46] and CarthaGene software version 1.0 [47]. The origins of the alleles (A or B genomes) were determined by comparison to the alleles coming from the diploid progenitors of the amphidiploid. Mapdisto software was used in a first step, for the linkage group determination and marker ordering within each linkage group. A minimum LOD of 4 and maximum recombination fraction of 0.3 were fixed for the linkage group determination using the "find groups" command. The order of the markers within each linkage group was estimated using the "order" command. The markers that had not been placed at LOD 4 were tried at decreasing LOD, down to a LOD of 2 and a maximum recombination fraction of 0.3. These markers are indicated in italic on the map (Figure 2). The quality of the genotyping data at a specific marker was controlled using the "drop locus" command. The few markers having bad quality genotyping data were discarded from the linkage analysis. In a second step, CarthaGene software was used for the optimization of the best marker order determined by Mapdisto. This was done applying the simulated "annealing" and "greedy" algorithms. The best maps obtained were improved using the "Flips" and the "Polish" commands. Genetic distances between markers were computed using Kosambi mapping function.

Introgression analysis

From the map of 298 SSR markers previously developed on the BC₁F₁ generation, a framework map comprising 115 SSR markers was derived. Compared to the initial map, this framework offered a regular coverage of all the linkage groups. These 115 SSR markers were used to genotype 123 BC₂F₁ individuals.

Introgression analysis of the BC₁F₁ and BC₂F₁ populations was performed using the CSSL Finder software version 0.8b4 [48]. To select a subset of BC₁F₁ and BC₂F₁ lines providing an optimal coverage of donor genome into the recurrent background, we imposed a target length of the introgressed wild segments of 20 cM, an overlapping of adjacent segments for a given LG and the best possible return to the background of the cultivated variety.

The percentage of wild genome in the BC₁F₁ and BC₂F₁ generations and its relative diminution between the two generations, the mean size of wild introgression segments per LG and per generation, as well as the distribution of the wild segment lengths were estimated using the genotyping data available for each generation. The analysis was

conducted on LGs longer than 75 cM. The lengths of the introgressed segments were calculated as the sum of consecutive intervals having a heterozygous genotype plus half the size of each flanking interval having a recurrent homozygous genotype.

Results

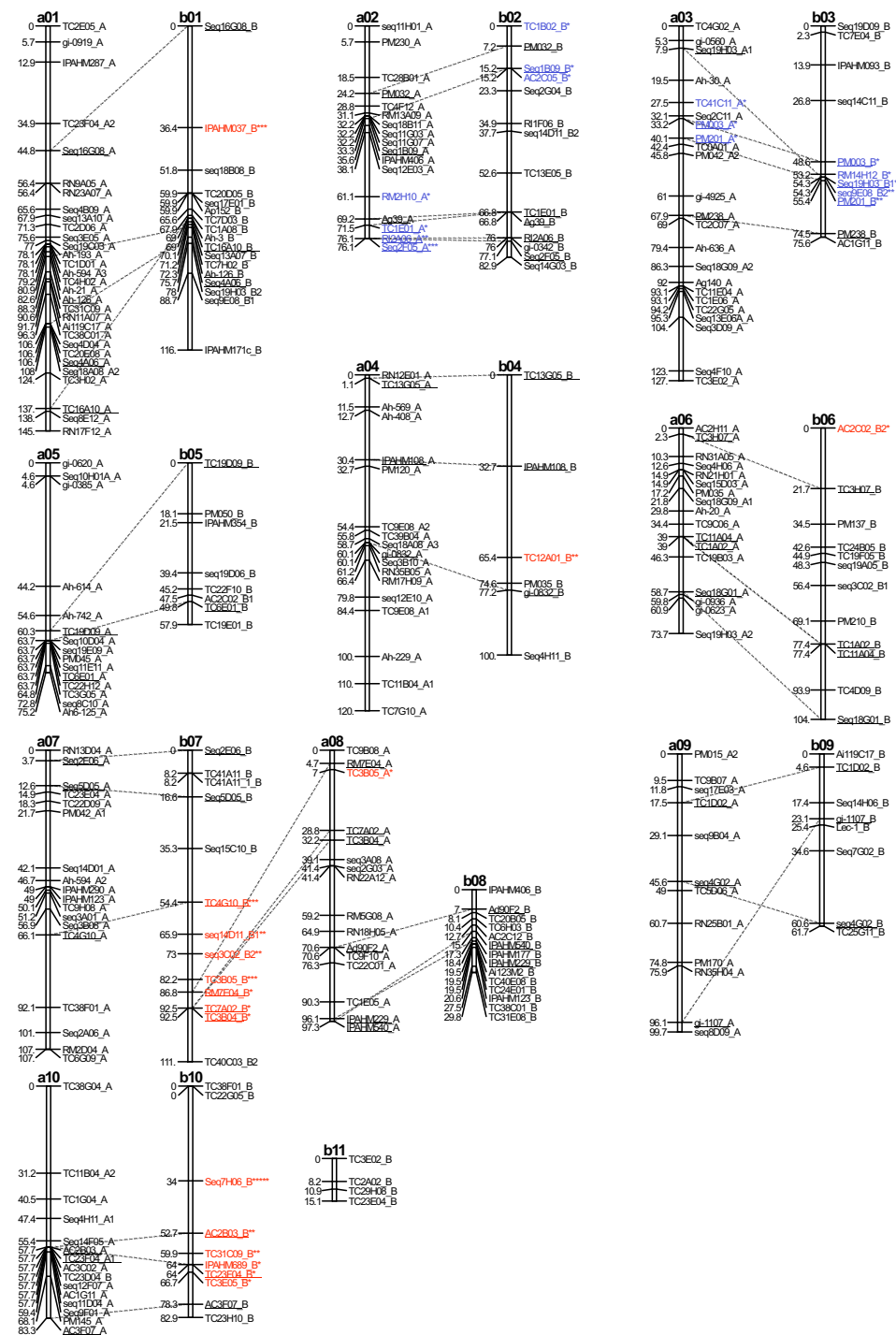
SSR polymorphism and origin of the markers

Among the 558 SSR markers screened, 333 (59.6%) were polymorphic between Fleur 11 and AiAd. At a given SSR locus, the sub-genomic origin of the alleles was determined by comparison with the alleles of the diploid parents of the amphidiploid *A. ipaënsis* and *A. duranensis* that were included on each gel. This allowed distinguishing three categories of markers among the 333 polymorphic markers: 174 SSRs that were polymorphic for the A genome (52.0%), 77 SSRs that were polymorphic for the B genome (23.0%) and 82 SSRs that were polymorphic for the two genomes (24.5%). The largest proportion of polymorphic markers originated from the A genome donor *A. duranensis* (76.6%), the B genome donor *A. ipaënsis* generating 47.6% of polymorphic markers.

Genetic map construction

Among the 333 polymorphic SSRs, we randomly selected 118 markers polymorphic for the A genome, all the markers polymorphic the B genome and those polymorphic for the two genomes. A total of 277 SSRs were used to genotype the population of 88 BC₁F₁ individuals. The 232 SSR markers that showed a clear electrophoretic profile amplified 322 loci. Finally, 298 loci were mapped in 21 linkage groups (LGs), spanning a total map distance of 1843.7 cM with an average distance of 6.1 cM between adjacent markers (Figure 2). The difference of polymorphism between the A and B genomes had an effect on the number of markers mapped on each genome, and the number and size of the linkage groups. For the A genome, 181 loci were mapped in 10 LGs with a number of markers per LGs varying between 12 and 30 (average of 18.1), and the length of the LGs ranging from 73.7 cM to 145.2 cM (average of 100.5 cM). For the B genome, 117 loci were mapped on 11 LGs with a number of markers per LGs varying between 4 and 17 (average of 10.7) and the length of the LGs ranging from 15.1 cM to 111.6 cM (average of 76.2 cM).

The comparison of the A and B genomes was undertaken using 53 SSR markers that mapped on both A and B LGs. The A and B LGs were considered to be homeologous when they shared at least 2 common markers. This allowed distinguishing 8 pairs of homeologous LGs (a01/b01, a02/b02, a03/b03, a04/b04, a05/b05, a06/b06, a09/b09 and a10/b10) and one quadruplet involving the LGs a07, b07, a08 and b08. LG a07 shared three markers with the upper part of LG b07 corresponding to at least the half

**Figure 2**

Genetic map and synteny between the A and the B genomes. The LGs deriving from the A genome are named from a01 to a10 and those deriving from the B genome from b01 to b11. Map distances are given in Kosambi centimorgans. Common markers between pair of homeologous LGs are underlined and connected with dashed lines. Markers placed at LOD < 4 are represented in italics, and those that amplified more than one locus on the same genome are identified by the number 1, 2 and 3. Loci showing significant segregation distortion ($P < 0.05$) are identified by stars following locus name. The colour and number of stars specify the direction and the intensity of the segregation distortion respectively. Blue: markers skewed toward the alleles of the cultivated parent. Red: markers skewed toward the alleles of the wild parent.

of this LG. The lower part of LG b07 shared three markers with the upper part of LG a08. Furthermore, the lower part of LG a08 shared three markers with LG b08 (Figure 2). The small LG b11 shared 1 marker with LG a03. An overall good collinearity was observed between homeologous LGs. However, three inversions of chromosomal segments were observed on the homeologous LGs a01/b01, a03/b03, and a09/b09. Small inversions were also observed on the homeologous LGs a08/b08. These inversions might result from artefacts as they concerned closely linked markers with more than one possible order having similar LOD values. No mosaic composition of linkage groups, where A genome markers would map together with B genome markers, was observed.

A total of 32 SSR markers (10.7%) showed significant segregation distortion at $P < 0.05$. Apart from 4 markers mapped on LGs b01 (IPAHM037), b04 (TC12A01), b06 (AC2C02), and LG a08 (TC3B05) all the distorted markers were concentrated in specific zones of 6 different LGs (a02, b02, a03, b03, b07 and b10). Differences between the A and B genomes were also observed. For the A genome, only 8 markers (4.4%) showed segregation distortions compared to 24 (20.3%) for the B genome. For the A genome, in the zones of distortion of LGs a02 and a03, all the distorted markers were skewed toward the alleles of the cultivated parent. For the B genome, the zones of distorted markers of LGs b02 and b07 were skewed toward the allele of the wild parent while those of b03 and b10 were skewed toward the allele of the cultivated parent.

Fifteen primer pairs (AC2C02, Ah-594, PM042, Seq14D11, Seq18A03, Seq18G09, Seq19H03, Seq3C02, Seq4H11, Seq9E08, TC11B04, TC9E08, TC19E01, TC23F04 and TC40C03) amplified consistently more than one locus on the same genome. We were able to map the duplicated loci for the markers Ah-594, PM042, Seq18A03, Seq18G09, Seq19H03, TC11B04 and TC9E08. Apart from the loci amplified by TC9E08 that mapped on the same LG (a04), the loci amplified by AC2C02, Ah-594, PM042, Seq14D11, Seq18A03, Seq18G09, Seq19H03, Seq3C02, Seq4H11, Seq9E08, TC11B04, TC19E01, TC23F04 and TC40C03 mapped on different LGs suggesting possible segmental duplications. These markers were identified by the number 1, 2 and 3 on the map (Figure 2).

Comparison with peanut published genetic maps

Conserved structural features between tetraploid maps

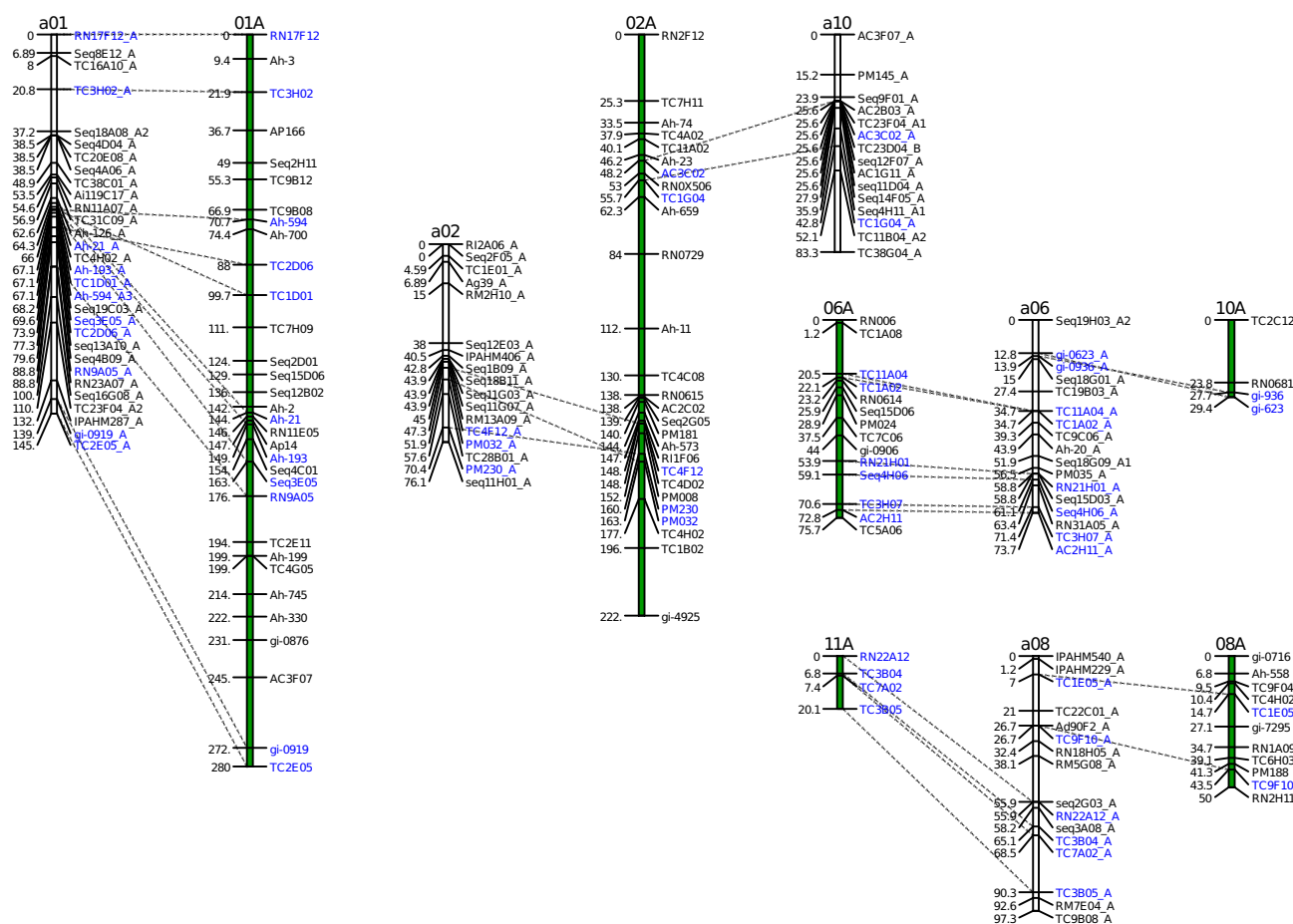
The present tetraploid map was compared to the RFLP based tetraploid BC₁F₁ map published by Burow et al. [18], further called "Burow's map" involving a cross between *A. hypogaea* variety Florunner and the synthetic amphidiploid TxAG6 [(*A. batizocoi* × (*A. cardenasii* × *A. diogoi*)]^{4X}. *A. batizocoi* was considered to be the B genome

donor and *A. cardenasii* and *A. diogoi* were the donors of the A genome. In that cross, 23 LGs spanning a total genetic distance of 2210 cM (Kosambi mapping function) were obtained. This map size was slightly larger than our map. A similar number of loci had been mapped on the A genome (156 for Burow's map versus 181 for our map) but the number of loci mapped on the B genome of the Burow's map was about 2 fold larger than on the present map (206 versus 117 respectively). The mean length of the A genome LGs was similar between the 2 maps (93.7 cM for the Burow's map vs 100.5 cM for our map) while the mean length of the B genome LGs of the Burow's map was 1.2 larger than that of the present map (94.1 cM for the Burow's map versus 76.2 cM for our map). The difference in map size between the two studies seems related to the difference of the number of mapped markers on the B genome. Interestingly, conservation of synteny between one B genome LG and two A genome LGs was observed in the two maps. On our map, LG b07 shared common markers with LG a07 and a08 while on Burow's map, LG 19 shared common markers with LG 9.1 and 9.2. Moreover, Burow et al. [18] has reported structural differences, mainly chromosome segment inversion, between four pairs of homeologous LGs (LG1/LG11, LG7/LG17, LG4/LG14 and LG5/LG15). Inversions of chromosomal segments have been observed for at least 3 LGs in our map (a01/b01, a03/b03 and a09/b09).

Comparison with diploid map

The results from the present tetraploid map were compared to the SSR based diploid F₂ map [15], involving a cross between two wild diploids with A genome, *A. duranensis* and *A. stenosperma*. In that population, 11 LGs covering a total map length of 1230.8 cM (Kosambi mapping function) have been described. The total map length was slightly longer than what we obtained in our map when considering the total size of the LGs of the A genome (1005.2 cM). The proportion of distorted markers found in the study of Moretzsohn et al. [15] was higher than what we recorded for our A genome map (50% versus 4%). Given that a similar number of individuals were used for the map construction in the two studies, the length difference between the 2 maps might be related to the higher proportion of distorted markers on the Moretzsohn's map.

The synteny between the 2 maps was assessed with 57 common SSR markers. For all the 10 LGs of the A genome of our map, we could identify corresponding LGs in the diploid map with an overall good collinearity. The salient features of the comparison of the two maps are shown in Figure 3. The number of common SSR markers per homologous LGs varied between 2 and 11. However the synteny was not conserved for four LGs of our map when compared to the diploid map. LG a02 and a10 of our map shared 3 (PM230, PM032 and TC4F12) and 2 (AC3C02

**Figure 3**

Salient features of the comparison between the A genome LGs of our tetraploid BCI map and the diploid AA map published by Moretzsohn et al. (2005). The LGs from this study are named a01, a02, a10, a06, and a08. The LGs of Moretzsohn's map (01A, 02A, 06A, 10A, 08A and 11A) are represented by a green bar. The LGs of Moretzsohn's map (01A, 02A, 06A, 10A, 08A and 11A) are represented by a green bar. Common markers between corresponding LGs in the two maps are indicated in blue, underlined and connected with dashed lines. In the two maps the distances are given in Kosambi centimorgans.

and TC1G04) markers with LG 2 of the diploid map respectively. LGs 8 and 11 of the diploid map shared 2 (TC1E05, TC9F10) and 4 (RN22A12, TC3B04, TC7A02 and TC3B05) markers with LG a08 of our map respectively. Moreover, LG a06 of our map that was homologous to LG 6 of the diploid map shared also 2 common markers (gi-936 and gi-623) with LG 10 of the diploid map. For LGs a06 and a08 of our map, there was no evidence of spurious linkage between two different LGs as all the markers in these LGs were mapped at LOD ≥ 4 .

Introgression analysis

In the BC₁F₁ generation, the percentage of heterozygous genome varied between 26.5% and 77.0% (average of 49.8%) while in the BC₂F₁ generation it varied between 6.1% and 44.4% (average of 22.2%). This percentage is

slightly inferior to the expected 25%, which is consistent with the selection that occurred at each generation for the best possible return to the background of the cultivated variety. From BC₁F₁ to BC₂F₁, we noted more than 50% reduction of the wild allele contribution to the genotypic constitution of the BC₂F₁ individuals. The distribution of the lengths of the wild segments in the BC₁F₁ and the BC₂F₁ generations was calculated for 14 LGs having a length comprised between 75 and 145.2 cM (Figure 4). The average lengths of the wild introgressed segments into the background of the cultivated were 51.8 cM in BC₁F₁ and 34.9 cM in BC₂F₁. From BC₁F₁ to BC₂F₁ generations, the segment lengths decreased of 33%. As shown in Figure 4, more than 15% of the BC₁F₁ lines and 20% of the BC₂F₁ lines had segment lengths comprised between 20 and 30 cM.

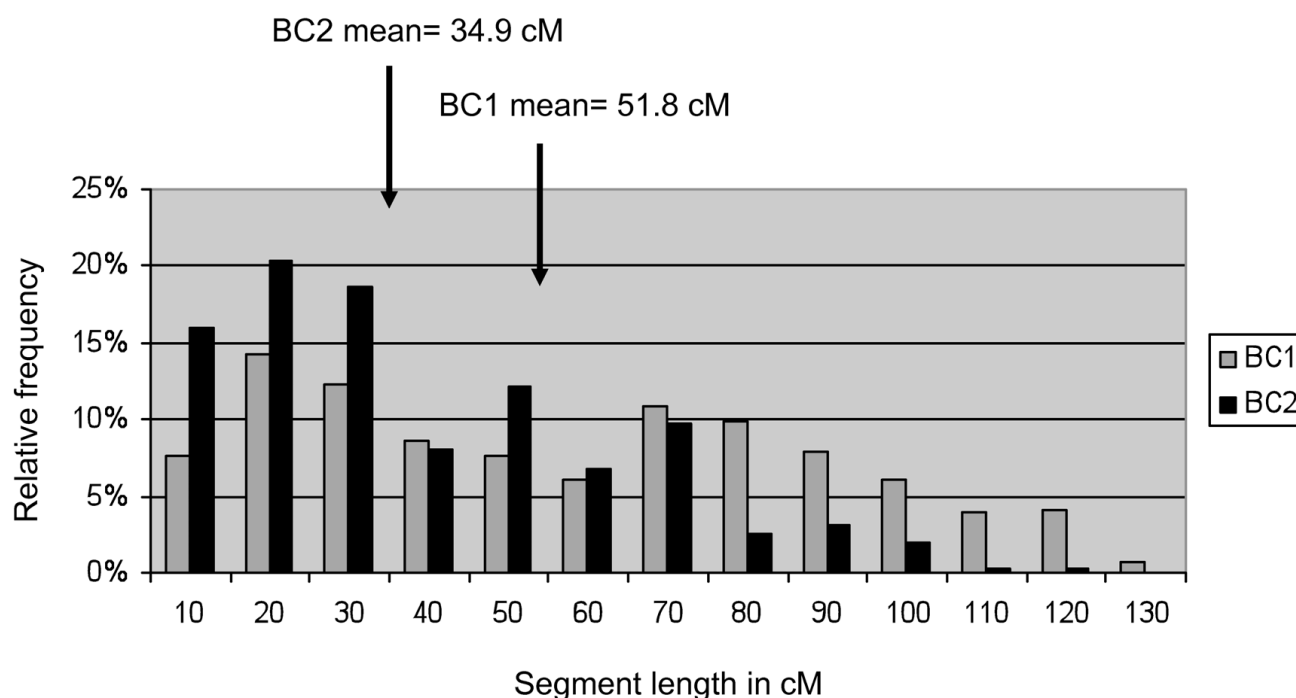


Figure 4
Distribution of donor segment lengths as calculated for both the BC₁F₁ and the BC₂F₂ generations derived from the cross between Fleur II cultivated variety and the synthetic amphidiploid AiAd.

The CSSL-Finder software was used to select a subset of BC₁F₁ lines and, then, of BC₂F₁ lines, which ensured, in each generation, an optimal coverage of the wild genome with overlapping target segment lengths of 20 cM between neighbouring lines and the best possible return to the cultivated background. In the BC₁F₁ population, a subset of 22 lines was selected. The segment lengths ranged between 2.3 cM and 89.4 cM (mean of 34.8 cM). All the adjacent segments were in overlapping position and the genome percentage of the recurrent cultivated variety ranged between 38% and 68% (mean 52%). In the BC₂F₁ population, a subset of 59 lines was selected. The segment lengths ranged between 2.3 cM and 46.9 cM (mean of 24.5 cM) and the percentage of the recurrent background between 62% and 94% (mean of 79%). A graphical representation of the BC₂F₁ selected lines is shown in Figure 5. The level of coverage of the wild introgressed segments in the background of the cultivated variety was optimal both in the BC₁F₁ and BC₂F₁ populations.

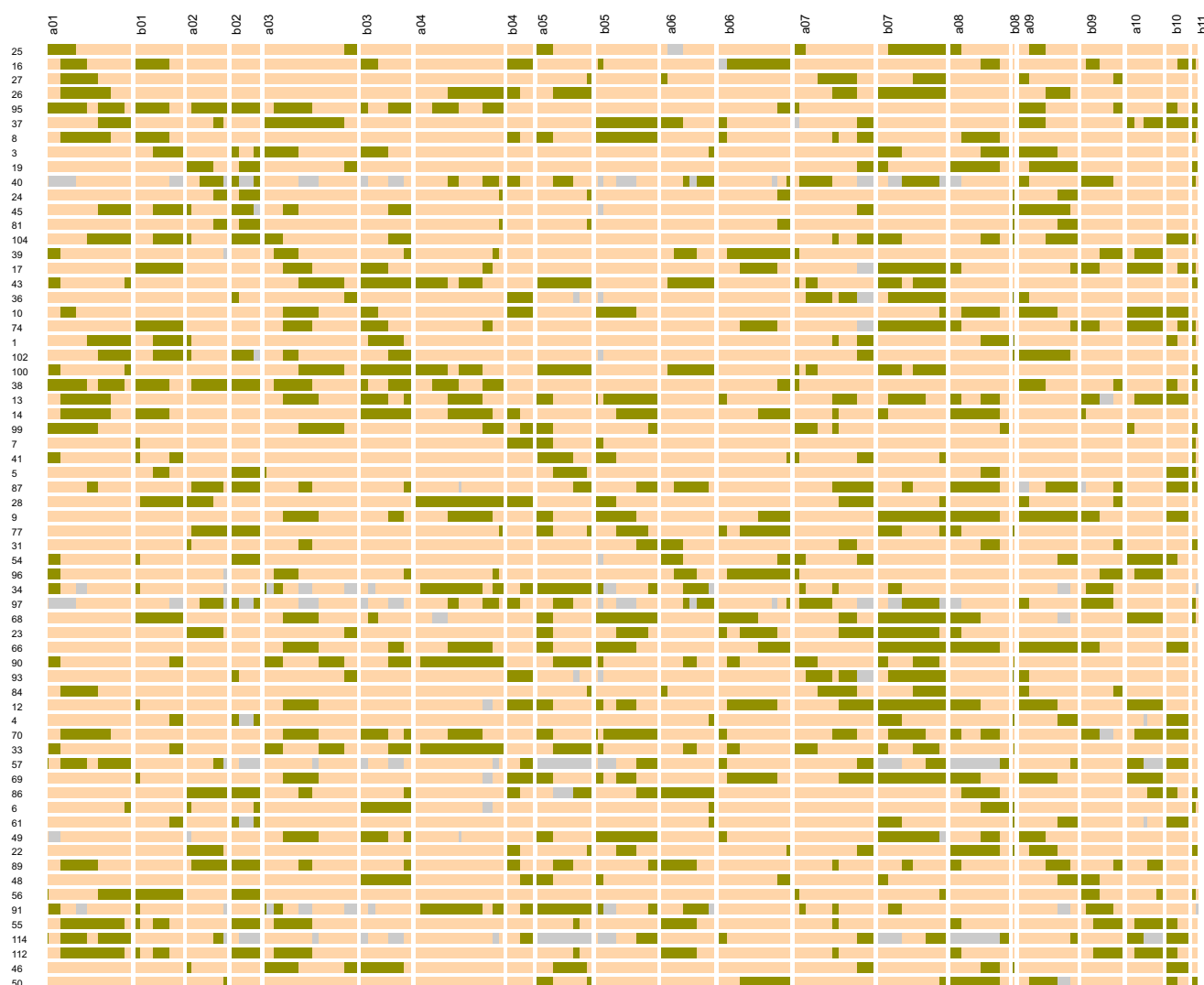
Discussion

In this study, the construction of a tetraploid molecular genetic map using a BC₁F₁ population and the development of a BC₂F₁ population allowed the analysis of the introgression of wild alleles in the background of a cultivated peanut variety. Several points have been highlighted

including (1) the low level of polymorphism of the SSR markers especially between the B wild genome *A. ipaënsis* and the B genome of the cultivated, (2) the collinearity between the A and B genomes, the synteny between tetraploid and diploid maps, and the similarity between tetraploid maps, (3) the good level of introgression of the wild genome segments in the background of the cultivated variety.

SSR polymorphism data is consistent with *A. duranensis* and *A. ipaënsis* being the most probable progenitors of the cultivated species

Cultivated peanut *Arachis hypogaea* is considered to be an allotetraploid ($2n = 4x = 40$) originated from a single hybridization event between two wild diploids with A and B genomes [3], followed by spontaneous duplication of the chromosomes. The identification of the wild progenitors of the cultivated peanut has been the object of numerous investigations using various approaches including cross-compatibility [5], molecular markers [4,12,20,21,49,50], biogeography [51], gene sequence comparison [52], physical mapping of rRNA genes [53] and Genome *In Situ* Hybridization (GISH) [6]. The most probable hypothesis gave *A. duranensis* as the wild donor of the A genome and *A. ipaënsis* as the wild donor of the B genome. In our study, a close relationship has been

**Figure 5**

Graphical genotype of the selected BC₂F₁ lines. Each row represented a candidate line and each column a Linkage Group. The green colour indicates the heterozygous (wild/cultivated) segments and the orange colour the homozygous regions for cultivated alleles. The gray colour indicates missing data.

observed between the putative wild progenitor's *A. duranensis* and *A. ipaënsis* and the cultivated *A. hypogaea* var. Fleur 11 based on 558 SSR markers. *A. duranensis* and *A. ipaënsis* shared 54.1% and 72.6% of common SSR alleles with the A genome and the B genome of *A. hypogaea* respectively. Moreover, 59.8% polymorphism was observed between the synthetic amphidiploid AiAd and the cultivated Fleur 11 variety. This is lower than the 83% of polymorphism that has been observed between the synthetic polyploid TxAG6 ($[(A. batizocoi \times (A. cardenasii \times A. diogoi)]^{4X})$) and the cultivated Florunner variety based on RFLP markers [18]. This result indicates that *A. duran-*

ensis and *A. ipaënsis* are more closely related to the A and the B genomes of the cultivated species than are *A. cardenasii* and *A. diogoi* for the A and *A. batizocoi* for the B genomes respectively. Moreover, the BC₁ tetraploid map obtained by crossing the synthetic amphidiploid AiAd and Fleur 11 indicated a disomic inheritance of all loci. For all the LGs obtained, the markers that were polymorphic for the A genome mapped on A LGs and those polymorphic for the B genome mapped on B LGs. The chromosome pairing seems to happen between "homologous genome" attesting the high affinity between *A. duranensis* and the A genome of the cultivated species, and

between *A. ipaënsis* and the B genome of the cultivated species. The same results have also been reported by Seijo et al. [6]. Our data fit well with the earlier reports indicating *A. duranensis* and *A. ipaënsis* as the most probable diploid progenitors of the cultivated peanut.

Genome rearrangements

In this study, the synteny analysis between the A and B genomes revealed inversion of chromosomal segments for at least three LGs, and a particular feature of synteny involving the LGs a07, b07, a08 and b08 (Figure 2). Conservation of synteny between the upper region of LGs a07 and b07, and between LG b08 and the lower region of LG a08 has been pointed out. Furthermore, the upper region of LG a08 shared also three markers with the lower region of LG a07 while LG b08 lacked a large chromosomal segment that could correspond to the region of conserved synteny between LGs b07 and a08. These observations are consistent with a major translocation event that has occurred between LGs b07 and b08. Similar feature of synteny conservation between two LGs of the B genome and two LGs of the A genome have also been reported at the diploid level when comparing the *A. duranensis* × *A. stenosperma* diploid AA map [15] and the *A. ipaënsis* × *A. magna* diploid BB map [16]. Interestingly, the quadruplet of syntenic LGs in the diploid maps was also found to be syntenic with those in our map (data not shown). These observations suggest that the rearrangement between LGs b07 and b08 is an ancient translocation event that happened prior to the tetraploidisation of the cultivated peanut.

Chromosome rearrangements, including the inversion of chromosomal segments within pairs of homeologous linkage groups and the conservation of synteny between a triplet of LGs (one LG of the B genome sharing common markers with two LGs of the A genome) have also been reported by Burow et al. [18]. We were not able to identify which LGs of our map are in synteny with those of the Burow's map due to the difference of the marker types used in the two studies.

The similarity of the rearrangement events observed in the diploid and the tetraploid maps, which involve different species for A and B genomes, suggests that these evolutionary mechanisms have contributed to the divergence of the A and B genomes of the section *Arachis*. It also raises the question of the classification of the B genome species. The relationships between species with B genome remain controversial. Using RFLP makers, Gimenes et al. [44] reported a clustering of *A. batizocoi* and *A. magna* which were less related to *A. ipaënsis*, while with SSR markers, Moretzsohn et al. [21] reported a clustering of the species with B genome including *A. batizocoi*, *A. magna* and *A.*

ipaënsis. Seijo et al. [6] reported, based on GISH, the distinction of *A. batizocoi* from the other B genome species and concluded that species with B genome do not seem to constitute a natural group.

The results obtained from the comparison of the diploid and tetraploid maps suggest that, based on the similarities of the rearrangement event, the species with B genome *A. ipaënsis*, *A. magna* and *A. batizocoi* could have derived from a common B genome ancestor and could be more closely related than what was previously reported based on molecular makers and on GISH. The construction of a consensus molecular genetic map involving the available diploid AA and BB maps and the tetraploid AABB maps as well as the study of crossability between species with B genome should shed new light on this issue.

Modifications of parental diploid genome following polyploidization, have been reported (for review see, [54,55]). The modifications include structural rearrangements, transposable element activation, difference in gene expression and epigenetic changes. These changes were observed in old polyploid [56-59] as well as in newly synthesized amphidiploids [60-62]. Rapid and dynamic changes in genome structure, including non additive inheritance of genomic fragments and genome-specific sequence deletion have been described in some taxa including *Brassica* [61] and wheat synthetic allotetraploids [63], but not in others including cotton [64] and sugarcane [65]. In peanut, Burow et al. [18] reported a possible genomic restructuring in the synthetic amphidiploid TxAG6 characterized by 5% of mapped alleles having unknown parental origins. In our study, we utilized a synthetic amphidiploid which had undergone several cycles of self-pollination before crossing with the cultivated allotetraploid. However, to the level of resolution afforded by our experiment, no change in genome structure has been pointed out. Further studies are needed to confirm the effectiveness and the level of genomic restructuring in peanut synthetic allotetraploid.

Wild segment introgressions and perspectives for the development of interspecific breeding populations

Few studies have been reported in the literature regarding the genetic mapping of introgressions from wild to the cultivated peanut. Apart from the study of Burow et al. [18], introgression mapping of wild segments in the background of a cultivated variety has been reported in 46 introgression lines originated from the hybridization between *A. cardenasii* × *A. hypogaea* [66]. Considering all the lines together, introgressed segments could be found on 10 of the 11 LGs of the *A. stenosperma* × *A. cardenasii* diploid AA map [14], and represented 30% of the diploid peanut genome. The mapping of a wild segment from *A.*

cardenasii conferring resistance to root-knot nematode [67] and the registration of two varieties of peanut 'COAN' [68] and 'NemaTAM' [69], having identifiable alleles conferring resistance to root-knot nematode transferred from wild species, have also been reported.

In our study, we used a synthetic wild amphidiploid as a mean for the introgression of alien alleles in the genetic background of a cultivated variety and, consequently, enlarging the genetic basis of the cultivated peanut. Genetic mapping of the wild introgressed segments gave a clear picture of the amount and the level of coverage of the wild donor genome in the background of the cultivated, and of the segment lengths and their relative decrease from BC₁F₁ to BC₂F₁ generation. The mean length of the wild segments was 51.8 cM in BC₁F₁ and 34.8 cM in BC₂F₁, and the decrease of segment size from BC₁F₁ to BC₂F₁ was about -33%. These values were similar to what was obtained in *Lycopersicon* wild × cultivated backcross populations [23]. Marker assisted selection of BC₁F₁ and then BC₂F₁ lines carrying the desirable donor segment with the best possible return to the background of the cultivated variety allowed the selection of a limited set of lines that offer an optimal coverage of the wild genome with an overlapping regions between neighbouring lines and an average segment lengths of 34.8 cM in BC₁F₁ and 24.5 cM in BC₂F₁, as well as a 79% return to the background of the cultivated variety in BC₂F₁. The rapid decrease of wild segment lengths observed between the BC₁F₁ and BC₂F₁ generations as well as the good level of recovery of the genetic background of the cultivated variety in BC₂F₁ generation is of great interest for the genetic mapping of QTLs and the development of Introgression Line (IL) libraries. ILs carrying small wild segments in a constant cultivated genetic background have the advantages of reducing epistatic and linkage drag effects and of improving the resolution of QTL mapping [23,70]. Furthermore, ILs are reliable and stable genetic resources that can be multiplied and evaluated in various environments. Many valuable sources of resistance to biotic stresses including resistance to *Cercospora* leafspot [71], to root-knot nematode [67,72], to Peanut Bud Necrosis virus (PBNV) [73], to late leaf spot disease [74] and sources of tolerance to abiotic stresses including tolerance to thermal stress [75] and to drought (Soraya Bertioli, Vincent Vadez personal communication) were identified in peanut wild relatives. These sources can be used for genetic improvement in peanut.

The BC₁F₁ and BC₂F₁ populations that we developed are excellent starting points for the development of new breeding populations such as Advanced Backcross (AB) and Chromosome Segment Substitution Lines (CSSL)

populations for analysis of the wild alleles contribution to the improvement of cultivated peanut varieties.

Conclusion

In this study, a nearly saturated genetic map has been developed from a cross between the synthetic amphidiploid AiAd and the cultivated Fleur 11 variety. This allowed the synteny analysis of the A and B genomes, the comparison with diploid and tetraploid maps and the analysis of the introgression segments from the most probable wild progenitors of the cultivated peanut into the background of the cultivated Fleur 11 variety. The results of this study confirmed the close relationship between the wild diploids *A. duranensis*, *A. ipaënsis* and the cultivated peanut and highlighted structural rearrangements, such as chromosomal segment inversions and a major translocation event, between the A and B genome species. Finally, we showed that the low level of polymorphism reported between cultivated peanut can be overcome by using the wild species. The material we have produced in this study should facilitate the development of AB and CSSL breeding populations for the identification and utilization of valuable genes from the largely untapped reservoir of useful alleles that remained in the wild peanut species.

Authors' contributions

DF designed and coordinated the study, was involved in genotyping data production, carried out data analyses and map construction and drafted the manuscript. HAT, IF and ON carried out crosses and population development. RR and MNS were involved in DNA extraction and genotyping data production. APF produced the synthetic amphidiploid AiAd. DJB and JCG were involved in the design of the study. BC contributed to editing of the manuscript and helped in data analysis. JFR conceived, designed and coordinated the study, was involved in map construction, and editing of the manuscript.

Acknowledgements

The authors gratefully acknowledge Angelique D'Hont, Jean-Marc Lacape and Nabila Yahiaoui for the critical review of the manuscript. This study was funded by the Cirad and the Generation Challenge Programme.

References

1. Krapovickas A, Gregory WC: **Taxonomía del género *Arachis* (Leguminosae)**. *Bonplandia* 1994, **8**:1-186.
2. Holbrook CC, Stalker HT: **Peanut breeding and genetic resources**. *Plant Breeding Reviews* 2003, **22**:297-355.
3. Kochert G, Stalker HT, Gimenes MA, Galgalo ML, Lopes CR, Moore K: **RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae)**. *Am J Bot* 1996, **83**(10):1282-1291.
4. Milla SR, Isleib TG, Stalker HT: **Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers**. *Genome* 2005, **48**(1):1-11.
5. Favero AP, Simpson CE, Valls JF, Vello NA: **Study of the evolution of cultivated peanut through crossability studies among *Ara-***

- chis ipaensis*, *A. duranensis*, and *A. hypogaea*. *Crop Sci* 2006, **46**(4):1546-1552.
6. Seijo G, Lavia GI, Fernandez A, Krapovickas A, Ducasse DA, Bertoli DJ, Moscone EA: **Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double GISH.** *Am J Bot* 2007, **94**(12):1963-1971.
 7. Ramsey J, Schemske DW: **Pathways, mechanisms, and rates of polyploid formation in flowering plants.** *Annual Review of Ecology and Systematics* 1998, **29**(1):467-501.
 8. Soltis PS, Soltis DE: **The role of genetic and genomic attributes in the success of polyploids.** *Proc Natl Acad Sci USA* 2000, **97**(13):7051-7057.
 9. Kochert G, Halward T, Branch WD, Simpson CE: **RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species.** *Theor Appl Genet* 1991, **81**(5):565-570.
 10. Halward T, Stalker T, LaRue E, Kochert G: **Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.).** *Plant Mol Biol* 1992, **18**(2):315-325.
 11. He G, Prakash CS: **Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.).** *Euphytica* 1997, **97**:143-149.
 12. Hopkins MS, Casa AM, Wang T, Mitchell SE, Dean RE, Kochert GD, Kresovich S: **Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut.** *Crop Science* 1999, **39**:1243-1247.
 13. Varshney R, Bertoli D, Moretzsohn M, Vadez V, Krishnamurthy L, Aruna R, Nigam S, Moss B, Seetha K, Ravi K, et al.: **The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.).** *Theor Appl Genet* 2008, **118**(4):729-739.
 14. Halward T, Stalker HT, Kochert G: **Development of an RFLP linkage map in diploid peanut species.** *Theor Appl Genet* 1993, **87**(3):379-384.
 15. Moretzsohn M, Leoi L, Proite K, Guimarães P, Leal-Bertoli S, Gimenes M, Martins W, Valls J, Grattapaglia D, Bertoli D: **A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae).** *Theor Appl Genet* 2005, **111**:1060-1071.
 16. Moretzsohn MC, Barbosa AV, Alves-Freitas DM, Teixeira C, Leal-Bertoli SC, Guimarães PM, Pereira RW, Lopes CR, Cavallari MM, Valls JF, et al.: **A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome.** *BMC Plant Biol* 2009, **9**:40.
 17. Bertoli D, Moretzsohn M, Madsen L, Sandal N, Leal-Bertoli S, Guimarães P, Hougaard B, Fredslund J, Schauser L, Nielsen A, et al.: **An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes.** *BMC Genomics* 2009, **10**(1):45-45.
 18. Burrow MD, Simpson CE, Starr JL, Paterson AH: **Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.). broadening the gene pool of a monophyletic polyploid species.** *Genetics* 2001, **159**(2):823-837.
 19. Raina SN, Mukai Y: **Genomic in situ hybridization in *Arachis* (Fabaceae) identifies the diploid wild progenitors of cultivated (*A. hypogaea*) and related wild (*A. monticola*) peanut species.** *Plant Systematics and Evolution* 1999, **214**(1):251-262.
 20. Gimenes MA, Lopes CR, Galgalo ML, Valls JF, Kochert G: **RFLP analysis of genetic variation in species of section *Arachis*, genus *Arachis* (Leguminosae).** *Euphytica* 2002, **123**:421-429.
 21. Moretzsohn M, Hopkins M, Mitchell S, Kresovich S, Valls J, Ferreira M: **Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome.** *BMC Plant Biol* 2004, **4**:11-11.
 22. Eshed Y, Zamir D: **An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL.** *Genetics* 1995, **141**(3):1147-1162.
 23. Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, Beck-Bunn T: **Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*.** *Theor Appl Genet* 1996, **92**(2):213-224.
 24. Fulton TM, Beck-Bunn T, Emmatty D, Eshed Y, Lopez J, Petiard V, Uhlig J, Zamir D, Tanksley SD: **QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species.** *Theor Appl Genet* 1997, **95**(5):881-894.
 25. Fridman E, Carrari F, Liu Y-S, Fernie AR, Zamir D: **Zooming in on a quantitative trait for tomato yield using interspecific introgressions.** *Science* 2004, **305**(5691):1786-1789.
 26. Lippman ZB, Semel Y, Zamir D: **An integrated view of quantitative trait variation using tomato interspecific introgression lines.** *Current Opinion in Genetics & Development* 2007, **17**(6):545-552.
 27. Wan XY, Su JM, Wang CC, Shen CM, Li JM, Wang HL, Jiang L, Liu SJ, Chen LM, Yasui H, Yoshimura A: **QTL detection for eating quality of cooked rice in a population of chromosome segment substitution lines.** *Theor Appl Genet* 2004, **110**(1):71-79.
 28. Wang YM, Dong ZY, Zhang ZJ, Lin XY, Shen Y, Zhou D, Liu B: **Extensive de novo genomic variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.).** *Genetics* 2005, **170**(4):1945-1956.
 29. Li JZ, Huang HQ, Heinrichs F, Ganai MW, Röder RS: **Analysis of QTLs for yield, yield components, and malting quality in a BC 3-DH population of spring barley.** *Theor Appl Genet* 2005, **110**(2):356-363.
 30. Xu JL, Lafitte HR, Gao YM, Fu BY, Torres R, Li ZK: **QTLs for drought escape and tolerance identified in a set of random introgression lines of rice.** *Theor Appl Genet* 2005, **111**(8):1642-1650.
 31. Zhang X, Zhou S, Fu Y, Su Z, Wang X, Sun C: **Identification of a drought tolerant introgression line derived from dongxiang common wild rice (*O. rufipogon* Griff.).** *Plant Mol Biol* 2006, **62**(1):247-259.
 32. Zhao XQ, Xu JL, Zhao M, Lafitte R, Zhu LH, Fu BY, Gao YM, Li ZK: **QTLs affecting morpho-physiological traits related to drought tolerance detected in overlapping introgression lines of rice (*Oryza sativa* L.).** *Plant Science* 2008, **174**(6):618-625.
 33. Liu S, Zhou R, Dong Y, Li P, Jia J: **Development, utilization of introgression lines using a synthetic wheat as donor.** *Theor Appl Genet* 2006, **112**(7):1360-1373.
 34. von Korff M, Wang H, Léon J, Pillen K: **AB-QTL analysis in spring barley: III. Identification of exotic alleles for the improvement of malting quality in spring barley (*H. vulgare* ssp. *spon-taneum*).** *Molecular Breeding* 2007, **21**(1):81-93.
 35. Schmalenbach I, Léon J, Pillen K: **Identification and verification of QTLs for agronomic traits using wild barley introgression lines.** *Theor Appl Genet* 2009, **118**(3):483-497.
 36. Risterucci AM, Grivet L, N'Goran JA, Pieretti I, Flament MH, Lanaud C: **A high-density linkage map of *Theobroma cacao* L.** *Theor Appl Genet* 2000, **101**(5):948-955.
 37. Palmieri DA, Hoshino AA, Bravo JP, Lopes CR, Gimenes MA: **Isolation and characterization of microsatellite loci from the forage species *Arachis pintoi* (Genus *Arachis*).** *Molecular Ecology Notes* 2002, **2**:551-553.
 38. He G, Meng R, Newman M, Gao G, Pittman RN, Prakash C: **Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.).** *BMC Plant Biol* 2003, **3**:3-3.
 39. Ferguson ME, Burrow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchell S: **Microsatellite identification and characterization in peanut (*A. hypogaea* L.).** *Theor Appl Genet* 2004, **108**:1064-1070.
 40. Palmieri DA, Bechara MD, Curi RA, Gimenes MA, Lopes CR: **Novel polymorphic microsatellite markers in section *Caulorrhizae* (*Arachis*, *Fabaceae*).** *Molecular Ecology Notes* 2005, **5**:77-79.
 41. Bravo JP, Hoshino AA, Angelici CMLCD, Lopes CR, Gimenes MA: **Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*.** *Genetics and Molecular Biology* 2006, **29**:516-524.
 42. Hoshino AA, Bravo JP, Angelici CMLCD, Barbosa AaVG, Lopes CR, Gimenes MA: **Heterologous microsatellite primer pairs informative for the whole genus *Arachis*.** *Genetics and Molecular Biology* 2006, **29**:665-675.
 43. Proite K, Leal-Bertoli SCM, Bertoli DJ, Moretzsohn MC, Silva FRd, Martins NF, Guimarães PM: **ESTs from a wild *Arachis* species for gene discovery and marker development.** *BMC Plant Biol* 2007, **7**:7-7.
 44. Gimenes M, Hoshino A, Barbosa A, Palmieri D, Lopes C: **Characterization and transferability of microsatellite markers of the**

- cultivated peanut (*Arachis hypogaea*). *BMC Plant Biol* 2007, **7**(9):.
45. Cuc L, Mace E, Crouch J, Quang V, Long T, Varshney R: **Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*)**. *BMC Plant Biol* 2008, **8**(1):55-55.
 46. Lorieux M: **MapDisto, A Free User-Friendly Program For Computing Genetic Maps**. *Plant and Animal Genome XV conference*. San Diego, CA 2007:P958.
 47. De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T: **CARTHAGENE: multipopulation integrated genetic and radiated hybrid mapping**. *Bioinformatics* 2005, **21**(8):1703-1704.
 48. **MapDisto Genetics Software: CSSL Finder** [<http://mapdisto.free.fr/CSSLFinder/>]
 49. Singh AK, Smartt J: **The genome donors of the groundnut/peanut (*Arachis hypogaea* L.) revisited**. *Genetic Resources and Crop Evolution* 1998, **45**(2):113-116.
 50. Raina SN, Rani V, Kojima T, Ogihara Y, Singh KP, Devarumath RM: **RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species**. *Genome* 2001, **44**:763-772.
 51. Ferguson ME, Jarvis A, Stalker HT, Williams DE, Guarino L, Valls JFM, Pittman RN, Simpson CE, Bramel PJ: **Biogeography of wild *Arachis* (Leguminosae): distribution and environmental characterisation**. *Biodiversity and Conservation* 2005, **14**(7):1777-1798.
 52. Jung S, Tate PL, Horn R, Kochert G, Moore K, Abbott AG: **The phylogenetic relationship of possible progenitors of the cultivated peanut**. *J Hered* 2003, **94**(4):334-340.
 53. Seijo JG, Lavia GI, Fernandez A, Krapovickas A, Ducasse D, Moscone EA: **Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae)**. *Am J Bot* 2004, **91**:1294-1303.
 54. Soltis DE, Soltis PS: **Polyploidy: recurrent formation and genome evolution**. *Trends in Ecology & Evolution* 1999, **14**(9):348-352.
 55. Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF: **Evolutionary genetics of genome merger and doubling in plants**. *Annual Review of Genetics* 2008, **42**(1):443-461.
 56. Wendel JF: **Genome evolution in polyploids**. *Plant Mol Biol* 2000, **42**(1):225-249.
 57. Rieseberg LH: **Chromosomal rearrangements and speciation**. *Trends in Ecology & Evolution* 2001, **16**(7):351-358.
 58. Liu B, Wendel JF: **Epigenetic phenomena and the evolution of plant allopolyploids**. *Molecular Phylogenetics and Evolution* 2003, **29**(3):365-379.
 59. Clarkson JJ, Lim KY, Kovarik A, Chase MW, Knapp S, Leitch AR: **Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae)**. *New Phytologist* 2005, **168**(1):241-252.
 60. Adams KL, Wendel JF: **Polyploidy and genome evolution in plants**. *Current Opinion in Plant Biology* 2005, **8**(2):135-141.
 61. Chen ZJ, Ni Z: **Mechanisms of genomic rearrangements and gene expression changes in plant polyploids**. *BioEssays* 2006, **28**(3):240-252.
 62. Nicolas SD, Mignon GL, Eber F, Coriton O, Monod H, Clouet V, Huteau V, Loutanlen A, Delourme R, Chalhoub B, et al.: **Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids**. *Genetics* 2007, **175**(2):487-503.
 63. Liu B, Vega JM, Feldman M: **Rapid genomic changes in newly synthesized amphiploids of *Triticum* and *Aegilops*. II. Changes in low-copy coding DNA sequences**. *Genome* 1998, **41**(4):535-542.
 64. Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF: **Polyploid formation in cotton is not accompanied by rapid genomic changes**. *Genome* 2001, **44**(3):321-330.
 65. Jannoo N, Grivet L, Chantret N, Garsmeur O, Glaszmann JC, Arruda P, D'Hont A: **Orthologous comparison in a gene-rich region among grasses reveals stability in the sugarcane polyploid genome**. *The Plant Journal* 2007, **50**(4):574-585.
 66. Garcia GM, Stalker HT, Kochert G: **Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers**. *Genome* 1995, **38**(1):166-176.
 67. Garcia GM, Stalker HT, Shroeder E, Kochert G: **Identification of RAPD, SCAR, and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea***. 1996:836-845.
 68. Simpson CE, Starr JL: **Registration of 'COAN' Peanut**. *Crop Science* 2001, **41**:918-918.
 69. Simpson CE, Starr JL, Church GT, Burrow MD, Paterson AH: **Registration of 'NemaTAM' Peanut**. *Crop Science* 2003, **43**:1561-1561.
 70. Zamir D: **Improving plant breeding with exotic genetic libraries**. *Nat Rev Genet* 2001, **2**(12):983-989.
 71. Stalker : **Utilizing *Arachis cardenasii* as a source of *Cercospora* leafspot resistance for peanut improvement**. *Euphytica* 1984, **33**(2):529-538.
 72. Sharma SB, Ansari MA, Varaprasad KS, Singh AK, Reddy LJ: **Resistance to *Meloidogyne javanica* in wild *Arachis* species**. *Genetic Resources and Crop Evolution* 1999, **46**(6):557-568.
 73. Reddy AS, Reddy LJ, Mallikarjuna N, Abdurahman MD, Reddy YV, Bramel PJ, Reddy DV: **Identification of resistance to Peanut bud necrosis virus (PBNV) in wild *Arachis* germplasm**. *Annals of Applied Biology* 2000, **137**:135-139.
 74. Pande S, Rao JN: **Resistance of wild *Arachis* species to late leaf spot and rust in greenhouse trials**. *Plant Disease* 2001, **85**:851-855.
 75. Nautiyal PC, Rajgopal K, Zala PV, Pujari DS, Basu M, Dhadhal BA, Nandre BM: **Evaluation of wild *Arachis* species for abiotic stress tolerance: I. Thermal stress and leaf water relations**. *Euphytica* 2008, **159**(1-2):43-57.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



**Chapitre III FOSTERED AND LEFT BEHIND ALLELES IN PEANUT
(*ARACHIS HYPOGAEA* L.): INTERSPECIFIC QTL MAPPING
REVEALS FOOTPRINTS OF DOMESTICATION AND USEFUL
NATURAL VARIATION FOR BREEDING.**

Fostered and left behind alleles in peanut (*Arachis hypogaea* L.): interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding.

Daniel Fonceka^{*}, Hodo-Abalo Tossim[§], Ronan Rivallan^{*}, Hélène Vignes^{*}, Issa Faye[§],
Ousmane Ndoye[§], Brigitte Courtois^{*}, Jean-François Rami^{*}

^{*} Cirad, UMR Développement et Amélioration des Plantes, Montpellier F-34398, France.

[§] ISRA/Ceraas, Route de Khombole, BP 3320, Thiès Escale, Senegal

Advanced backcross QTL in peanut

Key words: Peanut, Advanced Backcross QTL, wild, domestication, polyploidy.

Corresponding author: Jean-François Rami

Cirad, UMR Développement et Amélioration des Plantes, TA A96/3, Avenue Agropolis,
Montpellier F-34398, France.

Tel: +33 4 67 61 58 00

Fax: +33 4 67 61 56 05

rami@cirad.fr

ABSTRACT

Polyploidy has created a severe genetic bottleneck, especially for species of monophyletic origin. Cultivated peanut is an allotetraploid harbouring limited genetic diversity, likely resulting from the combined effects of its single origin and domestication. Peanut wild relatives represent an important source of novel alleles that could be used to broaden the genetic basis of the cultigen. Using an advanced backcross population developed with a synthetic amphidiploid as donor of wild alleles, under two water regimes, we conducted a detailed QTL study for several traits involved in peanut productivity and adaptation as well as in the domestication syndrome. A total of 95 QTLs were mapped in the two water treatments. About half of the QTL positive effects were associated with alleles of the wild parent and several QTLs involved in yield components were specific to the water-limited treatment. QTLs detected for the same trait mapped to non-homeologous genomic regions, suggesting differential control of the expression of genes in subgenomes as a consequence of polyploidization. The noteworthy clustering of QTLs for traits involved in seed and pod size and in plant and pod morphology is a telltale signature of domestication and suggests, as in many crops, that a small number of loci have contributed to peanut domestication. Our findings open new avenues for peanut improvement using wild relatives.

INTRODUCTION

Polyploidy means that two or more complete sets of chromosomes of the same (autopolyploid) or different (allopolyploid) genomes are present in the same nucleus. It is a prominent and significant process in plant evolution (WENDEL 2000; ADAMS and WENDEL 2005). Polyploidy has been considered important in conferring adaptive value to some cultivated species by increasing the allelic diversity, maintaining genome-wide heterozygosity and allowing the emergence of novel phenotypic variation (SOLTIS and SOLTIS 2000; PATERSON 2005; UDALL and WENDEL 2006; FAWCETT *et al.* 2009). The stages of polyploid formation usually include reproductive isolation from the progenitors, resulting in severe genetic bottlenecks. However, as most polyploid species have been formed recurrently from their wild progenitors (SOLTIS *et al.* 2009), a moderate level of polymorphism has been kept in polyploid plants. Peanut (*Arachis hypogaea* L.) is an allotetraploid ($2n=4x=40$) native from South America with an AB genome. In contrast to the recurrent formation of several polyploid species, the allopolyploid structure of cultivated peanut is likely derived from a single hybridization between two wild diploid species followed by chromosome doubling

(KOCHERT *et al.* 1996). Consequently, its monophyletic origin and domestication effects have greatly narrowed the genetic basis of the cultigen.

The peanut primary gene pool comprises elite breeding lines and landraces of the cultivated species *A. hypogaea* and *Arachis monticola*, a close related wild tetraploid species of cultivated peanut (DWIVEDI *et al.* 2007). Although a large amount of phenotypic variation is conspicuous in this gene pool, only a limited level of DNA polymorphism between genotypes has been observed (HE and PRAKASH 1997; HOPKINS *et al.* 1999; GIMENES *et al.* 2007; CUC *et al.* 2008; LIANG *et al.* 2009). This lack of polymorphism has greatly hampered the application of new molecular tools for the genetic analysis of peanut. Recent peanut genomics initiatives have led to the development of several EST and genomic DNA based SSR markers (HE *et al.* 2003; FERGUSON *et al.* 2004; MORETZSOHN *et al.* 2005, 2009; PROITE *et al.* 2007) and construction of the first cultivated x cultivated genetic maps (VARSHNEY *et al.* 2008; HONG *et al.* 2010). However, there has been little progress on the integration of molecular markers in intraspecific peanut breeding programmes despite the challenge to obtain new varieties with resistance to diseases and tolerance to abiotic stresses.

The secondary gene pool of cultivated peanut mainly comprises wild diploid species ($2n=2x=20$) and represents an important source of novel alleles that can be used to improve the cultigen. Extensive work has been done to characterize genetic relationships between species of this gene pool and cultivated peanut using molecular markers (LU and PICKERSGILL 1993; RAINA *et al.* 2001; GIMENES *et al.* 2002; SINGH *et al.* 2002; MORETZSOHN *et al.* 2004; TALLURY *et al.* 2005) and cytogenetics (SEJO *et al.* 2007; ROBLEDO *et al.* 2009; ROBLEDO and SEJO 2010). Several wild diploid species have been hypothesized as the possible ancestors of the cultivated species. Recent studies have proposed *A. duranensis* (A genome) and *A. ipaensis* (B genome) as the most probable wild progenitors (MILLA *et al.* 2005; SEJO *et al.* 2007; BUROW *et al.* 2009). FAVERO *et al.* (2006) produced a synthetic amphidiploid resulting from a cross between *A. ipaensis* and *A. duranensis* and doubling of the chromosome number. This amphidiploid has produced fertile hybrids when crossed with each of the botanical varieties of *A. hypogaea*. Furthermore, resistances to several diseases have been identified in the wild compartment (SHARMA *et al.* 1999; REDDY *et al.* 2000; PANDE and RAO 2001; DWIVEDI *et al.* 2002) and QTLs for disease resistance were recently mapped in a cross involving wild diploid species (LEAL-BERTIOLI *et al.* 2009). Introgression of disease resistance genes from the wild diploid species *A. cardenasii* into an elite peanut variety has also been reported (SIMPSON and STARR 2001; SIMPSON *et al.* 2003). However, the effective transfer of genes from wild species to cultivated species was reported to be labor intensive

(GARCIA *et al.* 2006) and the introgression of genes involved in the variation of complex traits such as yield and drought tolerance related traits has never been attempted. Hence, genetic variation existing in the wild compartment remains largely underexploited.

Wild relatives represent an important source of genes that has been successfully tapped to improve productivity and adaptation in various crops (TANKSLEY and MCCOUCH 1997; GUR and ZAMIR 2004; HAJJAR and HODGKIN 2007). TANKSLEY *et al.* (1996) have proposed an efficient advanced backcross-QTL (AB-QTL) approach to detect and map valuable QTLs and to simultaneously transfer them from wild to cultivated species. This approach has been widely adopted for mapping and introgressing QTLs involved in complex traits in several species (SWAMY and SARLA, 2008 for a review).

QTL mapping in crosses between crops and their wild progenitors is also a powerful mean for identifying genomic regions involved in morphological and physiological changes that distinguish crops from their wild relatives (PATERSON 2002). These morphological and physiological differences that have resulted from plant evolution under anthropogenic influences have been included in a generic term known as the “domestication syndrome” (HAMMER 1984; HARLAN 1992). Features of the domestication syndrome have been shared in almost all agronomically important domesticated species. The targets of domestication include the loss of mechanisms for seed dispersal and dormancy, changes in plant growth habits and increases in the size of harvested plant parts (KONINANGA *et al.* 1996; BURKE *et al.* 2002; GEPTS 2004). A tremendous amount of information is now available on the process underlying domestication in plants from both a genetic and an archeological standpoint (SMITH 2001; ZEDER *et al.* 2006; PURUGGANAN and FULLER 2009), and several QTLs and genes involved in this process have been identified (PICKERSGILL 2007 for a review).

Notwithstanding the identification of the two most probable wild progenitors of cultivated peanut, little is known about peanut evolution under domestication, and the genomic regions associated with domestication have never been reported. We recently published an SSR-based genetic map constructed using a BC₁ population derived from a cross between the amphidiploid (*A. ipaensis* x *A. duranensis*)^{x4} and a cultivated peanut variety, and an analysis of the genome-wide introgression of wild DNA fragments in the BC₂F₁ generation (FONCÉKA *et al.* 2009). As a follow-up to this study, we have produced an AB-QTL population that represents a unique opportunity to map QTLs involved in peanut domestication and to explore the reservoir of agronomically interesting alleles remaining in the wild species. As peanut is mainly grown under rainfed conditions in the arid and semi-arid tropics, drought is one of the major factors limiting its production (REDDY *et al.* 2003; CLAVEL *et al.* 2005). An important

breeding objective is thus to develop varieties that can produce suitable yields under drought conditions.

In this article, we present a detailed QTL analysis of several traits involved in peanut productivity and adaptation under two water regimes as well as in the domestication syndrome. Based on these results, we report the identification of wild alleles that contribute positive variations to complex traits, we outline several regions of the peanut genome involved in the domestication process and we compare the subgenome contribution to trait variations.

MATERIALS AND METHODS

Population development

A population of 142 individuals (87 BC₃F₁ and 55 BC₂F₂) was produced using 44 BC₂F₁ plants derived from the cross between the cultivated Fleur11 variety, used as recurrent parent, and the amphidiploid AiAd (*A. ipaensis* x *A. duranensis*)^{x4} (FAVERO *et al.* 2006), used as donor parent. Fleur11, a local peanut variety grown in Senegal, is a Spanish type with an erect growth habit, low to moderate pod constriction, short cycle (90 days), high yielding, and tolerance to drought. The population was produced in greenhouse conditions at the Centre d'Etude Regional pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS), Thiès, Senegal. The breeding scheme for producing BC₂F₁ individuals has been described previously (FONCÉKA *et al.* 2009). Each of the 44 BC₂F₁ individuals used as female parent was: i) backcrossed with Fleur11 and ii) allowed to self-pollinate. A total of 565 seeds were harvested and sown individually in large deep pots in the greenhouse. DNA was extracted from young seedlings and BC₃F₁ individuals were differentiated from BC₂F₂ individuals using segregation analysis of 147 microsatellite markers. These markers were chosen to offer regular coverage of the genetic map produced previously (FONCÉKA *et al.* 2009). All BC₃F₁ and BC₂F₂ individuals were allowed to self-pollinate to produce BC₃F₂ and BC₂F₃ families that were then used for the phenotyping experiment. The choice of the 142 individuals (i.e. 87 BC₃F₁ and 55 BC₂F₂) retained as the final population was based on two criteria: i) maximization of donor allele frequencies in heterozygous or homozygous situations at each of the 147 SSR markers, and (ii) the number of seeds produced per BC₃F₁ and BC₂F₂ individual, which can be a strong limitation factor in peanut.

Field preparation

The experiment was conducted in the field from September to December 2009 at the Centre National de Recherche Agronomique (CNRA) in Bambey (14.42° N and 16.28W°), Senegal. In this research station, the soil was ferruginous, with 90% sand content, and low clay content (3–6%). Forty-five days before sowing, the field was plowed to eliminate weeds. One hundred and fifty kg/ha of organic fertilizer and 1 t/ha of mineral fertilizer (6-20-10) were added 4 weeks and 3 days before sowing, respectively. The field was kept manually weed-free before sowing and throughout the experiment.

Experimental design

A total of 144 plots (the 142 BC families and the Fleur11 parent repeated twice) were tested under two water regimes: well-watered and water-limited. For each water regime, an alpha experimental design was used with two replications and nine blocks per replication. The blocks contained 16 rows, 3 m each. The individuals were arranged in rows of 10 plants. The spacing was 30 cm between plants and 50 cm between rows. Due to the limited number of seeds per BC family, one seed per hill was sown at 4 cm depth. Before sowing, the seeds were treated with Granox (captafol 10%, benomyl 10%, carbofuran 20%) to protect them from insects and diseases.

Water management

In the geographical area of the Bambey research station, the rainy season lasts about 3 months, from early July to late September. The experiment was sown on 16th September 2009. In both treatments (well-watered, stressed), the total amount of water (rainfall + additional irrigation) received from the sowing date to 43 days after sowing (DAS) was 184 mm. After this date, corresponding to the pod filling stage, stress was applied in the stressed treatment by withholding irrigation until 84 DAS, representing a total stress duration of 40 days. Irrigation was restarted from 84 DAS to harvest (95 DAS) and 21 mm of water was added. The total amount of water received in the stressed treatment was 205 mm. In the well-watered treatment, irrigation was continued throughout the experiment until harvest, with 315 mm of water applied overall.

Soil moisture status

The soil volumetric water content was measured every 4 days to a depth of 1 m at 10 cm intervals, from the sowing date to the end of the stress (84 DAS) in both treatments using a Diviner 2000 capacitance profile probe (Sentek Environmental Technologies, Stepney,

Australia). In each treatment, the measurements of 12 access tubes were averaged. The access tubes were randomly scattered in the treatment plots. Variations in the soil water availability at each depth from 10 cm to 1 m was expressed as a fraction of transpirable soil water (FTSW) using the following formula:

$$FTSW (\%) = (SWC_{ERD} - SWC_{PWP}) / (SWC_{FC} - SWC_{PWP})$$

where SWC_{ERD} , SWC_{PWP} and SWC_{FC} correspond to the soil water content at effective rooting depth, permanent wilting point, and field capacity, respectively (CLAVEL *et al.* 2004). The field capacity value was obtained from a previous study carried out in Bambey (SARR *et al.* 2004).

Trait evaluation

Except for the days to flowering, the plant growth habit and the main stem height, all traits were recorded after harvest. The plants were harvested row by row 95 days after sowing and exposed to ambient temperature (30-35°C) for 1 month to allow complete drying of haulms and pods. A total of 27 traits were evaluated in BC_3F_2 and BC_2F_3 families under one or two water regimes. The phenotypic value of each trait for each BC_3F_1 and BC_2F_2 individual was obtained by averaging the values of the corresponding BC_3F_2 and BC_2F_3 families. The trait values were then expressed on a per plant basis. As indicated above, due to the limitation in the seed number, one seed per hill was sown. The traits were thus evaluated on a minimum of 3 to a maximum of 10 plants (mean 8.5 plants) per family and per replication. Details on the trait measurements are given below.

Days to flowering

The number of days from sowing to flowering (**DFL**) was evaluated on the basis of the flower appearance.

Plant architecture

The plant growth habit (**GH**) and main stem height (**PH**) were recorded in the well-watered conditions at the podding stage and 60 days after sowing, respectively. The plant growth habit was recorded on a 1-6 scale using the descriptors for groundnut (IBPGR and ICRISAT 1992) where, 1= procumbent 1, 2=procumbent 2, 3= decumbent 1, 4= decumbent 2, 5=decumbent 3,

and 6= erect. The main stem height was measured from the cotyledonary axil up to the terminal bud.

Pod morphology

Pod beak (**PB**), constriction (**PC**), length (**PL**) and width (**PWI**) were evaluated on 30 pods. All traits were measured in the two water regimes, except for pod constriction, which was measured in the well-watered condition only. Pod beak and constriction were recorded on a 0-9 scale according to the descriptors for groundnut. Pod length and width were measured using a caliper with a digital display.

Seed morphology

Seed length (**SL**) and width (**SWI**) were evaluated on 30 seeds in the two water regimes using a caliper with a digital display.

Yield components.

The yield components were determined in the two water regimes based on the pod, haulm and seed dry mass. The total biomass was first weighed to determine the total biomass per plant (**TB**). The pods were removed and weighed to determine the total pod weight per plant (**PW**). The haulm weight per plant (**HW**) was calculated as the difference between the total biomass and the total pod weight. The total number of pods per plant (**PN**) was determined and 100 pods were randomly sampled and weighed (**HPW**). The 100 pods were shelled and mature pods with a dark internal pericarp color were counted (**PMAT**). All seeds contained in the 100 pods were weighed and mature seeds were separated from immature seeds and counted. The weight of 100 seeds (**HSW**) was calculated as the weight of mature seeds divided by the number of mature seeds multiplied by 100. The total seed weight per plant (**SW**) and the total number of mature seeds per plant (**SN**) were estimated based on the total number of pods (**PN**), the number of mature seeds in 100 pods (**NMSH**) and the weight of seeds in 100 pods (**SWH**):

$$SW = (SWH/100) * PN \text{ and } SN = (NMSH/100) * PN.$$

The shell weight (**SHW**) was computed as the difference between the 100 pod weight and the 100 seed weight relative to the total number of pods per plant:

$$SHW = ((HPW - HSW)/100) * PN$$

The harvest index (**HI**) was calculated as a percentage of the pod weight to the total biomass.

Stress tolerance indices (STI) were calculated for the pod number (**STI-PN**), seed number (**STI-SN**), 100 pod weight (**STI-HPW**), 100 seed weight (**STI-HSW**), pod weight (**STI-PW**), seed weight (**STI-SW**), haulm weight (**STI-HW**) and total biomass (**STI-TB**) using the following formula:

$$STI = (Y_{is} * Y_{iw}) / (\bar{Y}_m)^2$$

with Y_{is} = phenotype value of individual (i) for a given trait in the stressed condition, Y_{iw} = phenotype value individual (i) for a given trait in the well-watered condition and \bar{Y}_m = trait mean of all genotypes in the well-watered condition (FERNANDEZ 1992).

Statistical analysis

Qualitative data such as the plant growth habit (GH), pod constriction (Pc) and beak (Pb) were first transformed to quantitative data using the ratio of the frequency of a phenotypic class by the total number of observations. All statistical analyses were performed using the R statistical programming language (R DEVELOPMENT CORE TEAM 2010). Basic statistical analyses (mean and standard deviation) were calculated for each trait. The data normality was checked with the Shapiro test for normality. An analysis of variance (ANOVA) was performed to estimate the genetic and replication effects on each trait under each water treatment. This was done following a standard linear model with genotype, replication, block and interaction effects, as

$$Y_{ijk} = \mu + G_i + r_j + b_{jk} + e_{ijk}$$

with Y_{ijk} = observed value for a given trait, μ = mean of the population, G_i = genotype effect, r_j = replication effect, b_{jk} = block within replication effect and e_{ijk} = residual error. In addition, a combined analysis of variance for the two water regimes was performed following a standard procedure of a fixed model with genotype, water regime, replication, and block and interaction effects, as

$$Y_{ijkl} = \mu + G_i + W_j + r_{jk} + b_{jkl} + G_i * W_j + e_{ijkl}$$

with Y_{ijkl} = observed value for a given trait, μ = mean of the population, G_i = genotype effect, W_j = water regime effect, r_{jk} = replication within water regime effect, b_{jkl} = block within replication and water regime effect, $G_i * W_j$ = genotype x water regime interaction and e_{ijkl} = residual error.

Estimates of broad-sense heritability were calculated as

$$h^2_b = \sigma^2_G / (\sigma^2_G + \sigma^2_E) \text{ with } \sigma^2_G = (MS_G - MS_E) / r \text{ and } \sigma^2_E = MS_E$$

where σ^2_G is the genotypic variation, σ^2_E the residual variation, MS_G and MS_E the genetic and residual mean squares and r the number of replications.

Data for each water regime were analysed using a linear mixed model fitted with the R/lme4 software package. In the model, we considered replications and blocks within replications as fixed effects and genotypes as random effects. Best linear unbiased predictors (BLUP) were extracted from this model for each genotype and trait and used for the QTL analyses.

Molecular analysis and QTL identification

The genetic map that had been previously produced (FONCÉKA *et al.* 2009) was used for the QTL analysis. In accordance with their position on the genetic map 147 microsatellite markers were used to genotype the 142 individuals of the population. Interspecific advanced backcross populations (BC₂F₂, BC₃) carry the risk of a low frequency of the wild donor allele at some loci. To overcome this situation, we checked the genotypic composition of the population at each marker and assessed the number of individuals in each genotypic class (i.e. homozygous for the recurrent parent, heterozygous and homozygous for the donor parent). For some loci, the number of individuals homozygous for the donor parent was below 5. Genotypes at these markers were replaced by missing data. For the QTL identification, standard interval mapping (SIM) was performed with the Haley-Knott regression method using the R/qtl package (BROMAN *et al.* 2003). Individuals in generation BC₃F₁ and BC₂F₂ were considered separately to calculate the genotypic probabilities at each 1 cM interval using MDM and GRAFGEN software (SERVIN *et al.* 2002). The genotypic probabilities of both generations were then combined and the resulting file was considered as a unique F₂ population and further imported into R/qtl. QTL detection for each trait and treatment, was performed using the following model:

$$y = \mu + \beta x + \alpha z_1 + \delta z_2 + \varepsilon$$

where, y is the observed phenotype, μ the mean of the population, α and δ the additive and dominance effects of the putative QTL, respectively, z_1 and z_2 are the probabilities for QTL genotypes conditional to the flanking marker genotypes, x an additive covariate ranging from 1 to 44 representing the BC₂ family, β the BC₂ family covariate effect, and ε the residual error.

A two-dimensional two QTL genome scan method was also used to test for the presence of two QTLs in the same linkage group. This was applied in some particular cases when the LOD curve of the single QTL genome scan method displayed two distinct LOD peaks for a given linkage group. A LOD threshold value to indicate a significant QTL effect was determined for each trait using 1000 permutations with a genome-wise significance level of $\alpha = 5\%$ (CHURCHILL AND DOERGE 1994). The confidence interval estimates of the QTL location

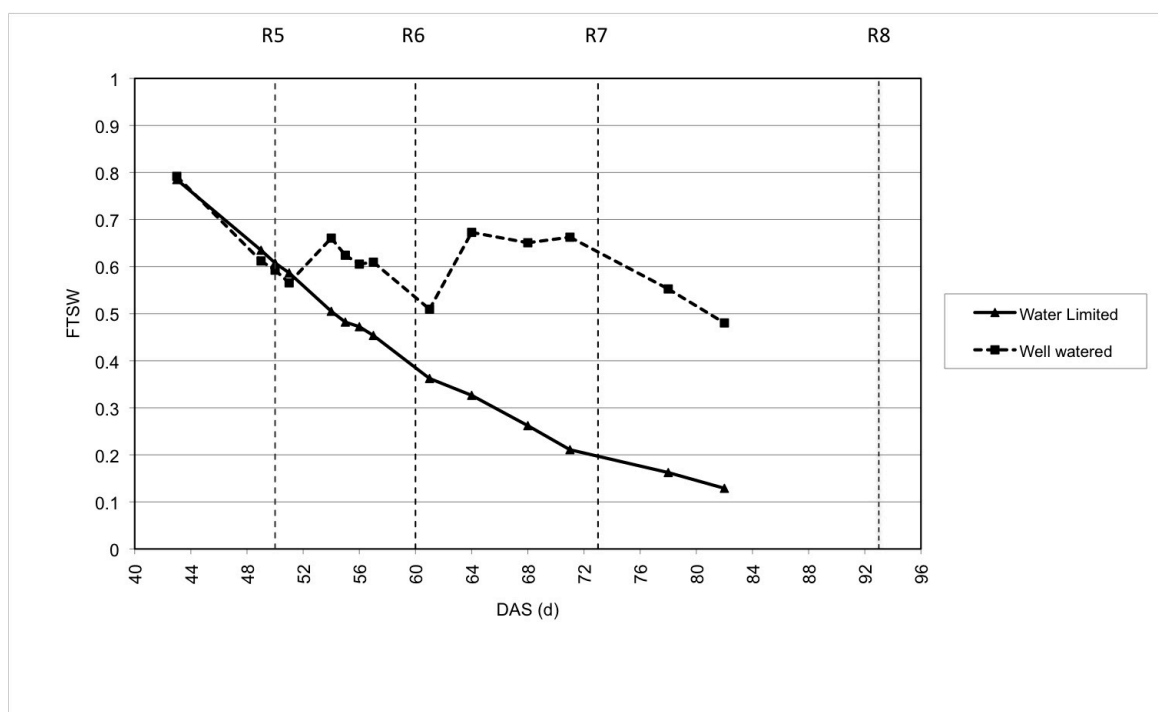


Figure 1: FTSW variation during population evaluation in well watered and water limited treatments. R5 to R8 correspond to peanut reproductive stages.

was obtained using the 1.5-LOD support interval method (BROMAN and SEN 2009). The proportion of phenotypic variance (R^2) explained by each QTL was obtained by fitting a model including the QTL and the BC2 family covariate. The proportion of phenotypic variance (R^2) explained by all detected QTLs for a given trait was obtained by fitting a model including all detected QTLs and the BC2 family covariate. The graphical representation of the QTLs was obtained using Spidermap software (Rami, unpublished).

RESULTS

Soil moisture status and stress intensity

The soil moisture measurement results showed that the calculated FTSW values were maintained between 0.80 and 0.50 in the well-watered treatment (Figure 1). This range of FTSW is generally considered sufficient to keep plants out of stress (SADRAS and MILROY 1996). Based on the FTSW values, we divided the stress duration for the water-limited treatment into three different periods corresponding to different stress intensity levels. A period of 15 days, from 43 DAS (date of irrigation interruption) to 58 DAS, during which the FTSW values were between 0.80 and 0.45, corresponded to an absence of stress, a period of 12 days from 59 DAS to 71 DAS during which the FTSW values decreased from 0.45 to 0.20, corresponded to a moderate stress intensity, and a period of 12 days from 72 DAS to 84 DAS, with FTSW values under 0.20, corresponded to a severe stress intensity. The two latter stress intensity periods were characterized by the wilting of plants in the afternoon for the moderate stress conditions, and at mid-morning for the severe stress conditions. These three phases roughly corresponded to peanut reproductive stages R5, R6-R7 and R7-R8, respectively, as described by BOOTE (1982). The different pod developmental stages are indicated in Figure 1. Moderate stress occurred from the end of seed formation to the beginning of pod maturity (R6-R7) and severe stress occurred from the beginning of pod maturity to harvest maturity (R7-R8).

Trait variability and heritability

The phenotypic values were normally distributed for most of traits in both conditions (Figure S1). The population mean for each trait in each condition tended to be skewed towards the phenotypic value of the recurrent Fleur11 parent (Table 1). A small range of variation (4 days) was observed for DFL, with values ranging from 18 to 21 days after sowing. Conversely, a high level of variability was observed for the morphological traits. The plant growth habit (GH) showed a wide range of morphologies, ranging from completely prostrate

Table 1: List of traits and descriptive statistics in well-watered and water limited treatments (Fleur11: value measured for the cultivated parent, SD: standard deviation, h^2 : heritability, Stn: Shapiro-Wilk test for normality)

Trait category	Trait name	Symbol	Unit	Well-watered					Water limited				
				Mean	Fleur11	SD	h^2	Stn	Mean	Fleur11	SD	h^2	Stn
Flowering	<i>Days to flowering</i>	DFL	days	19.3	19.3	0.78	0.77	n	-	-	-	-	-
Plant architecture	<i>Growth habit</i>	GH	1-6 scale	5.2	6.0	0.62	0.81	n	-	-	-	-	-
	<i>Main stem height</i>	PH	cm	20.3	22.4	2.52	0.86	y	-	-	-	-	-
Pod morphology	<i>Pod width</i>	PWI	mm	10.9	11.3	0.66	0.74	n	12.6	12.7	0.8	0.87	n
	<i>Pod length</i>	PL	mm	26.8	27.7	1.73	0.71	y	28.5	28.0	1.7	0.86	y
	<i>Pod beak</i>	PB	0-9 scale	5.3	5.1	0.45	0.54	y	5.3	4.9	0.7	0.64	y
	<i>Pod constriction</i>	PC	0-9 scale	5.7	5.5	0.85	0.80	n	-	-	-	-	-
Seed morphology	<i>Seed length</i>	SL	mm	14.0	13.8	0.69	0.84	y	14.6	14.4	0.6	0.81	n
	<i>Seed width</i>	SWI	mm	8.6	8.8	0.45	0.82	n	9.2	9.2	0.4	0.81	n
Yield components	<i>Total biomass</i>	TB	g	94.6	98.6	10.37	0.54	n	78.1	75.7	6.8	0.50	y
	<i>Pod weight</i>	PW	g	19.6	19.6	2.66	0.39	y	17.1	15.3	3.4	0.57	y
	<i>Haulm weight</i>	HW	g	75.1	79.2	8.8	0.55	n	60.9	60.1	4.5	0.46	n
	<i>Seed weight</i>	SW	g	14.3	14.5	1.89	0.39	y	11.8	10.8	2.7	0.62	y
	<i>Shell weight</i>	SHW	g	5.3	5.2	0.97	0.48	y	5.9	5.2	0.8	0.50	y
	<i>Hundred pod weight</i>	HPW	g	108.5	117.0	12.5	0.80	n	98.8	108.9	13.9	0.79	n
	<i>Hundred seed weight</i>	HSW	g	54.6	60.2	5.34	0.77	n	53.3	54.7	6.3	0.90	n
	<i>Pod number</i>	PN	count	18.8	17.9	2.5	0.44	y	17.5	15.4	1.9	0.53	y
	<i>Seed number</i>	SN	count	34.3	32.5	6.57	0.48	n	32.2	29.9	4.3	0.51	n
	<i>Harvest index</i>	HI	%	22.8	22.3	2.18	0.39	y	22.5	21.6	1.7	0.44	n
	<i>Percentage of maturity</i>	PMAT	%	0.7	0.7	0.027	0.41	y	0.5	0.5	0.0	0.39	n

to totally erect. A similar range of variation was observed for PB and PC, i.e. from inexistent to very prominent and from inexistent to very deep, respectively. A wide range of variation was also observed for the yield component traits (Table 1). In general, genotypes that outperformed the recurrent parent in terms of pod and seed number had smaller pods and seeds. However, we observed some genotypes that had a better performance than the cultivated parent in terms of number of pods, while keeping a similar 100 seed weight. Transgressive segregation was not observed for HSW, for which the best genotypes were similar to the recurrent parent.

The comparison between the two water treatments showed that water stress had a negative impact on the R6-R8 developmental stages, corresponding to grain filling and pod maturity. SW and PMAT were consequently the most affected traits, with a population average reduction of 21.0% and 18.9%, respectively. Water stress also negatively affected TB, PW and HW, with reductions of 17.4%, 12.8% and 17.2%, respectively, and to a lesser extent HPW, HSW, PN and SN, with reductions of 8.9%, 2.4%, 6.8% and 6.0%, respectively. The harvest index values were similar in the two conditions.

The analyses of variance showed significant differences ($P \leq 0.001$) between genotypes for each trait in the two water treatments. The estimated broad sense heritability values were similar in the two water treatments except for PW, HPW and SW for which they were higher in the water-limited treatment (Table 1). For TB, PW, HW, PN, SN, SHW, HI and PMAT, the heritability estimates were low, ranging from 0.39 to 0.54 in the well-watered treatment and from 0.39 to 0.62 in the water-limited treatment. Higher heritability estimate values were obtained for HPW, HSW, DFL and plant, pod and seed morphological traits ranged from 0.54 to 0.90 in both conditions (Table 1). The combined analysis of variance showed a significant genotype x water treatment interaction ($P \leq 0.05$) for a few traits, including PW, SN, HPW and PMAT (data not shown).

Correlation between traits

The phenotypic correlations between traits in the well-watered and water-limited treatments are shown in Table 2. The same trend was observed in the two water treatments. The highest values (up to 0.80) were obtained between TB and HW, between PW, PN, SN and SHW, and between HSW and HPW. HI was negatively correlated with HW but positively correlated with PW, PN, SW, SN and PMAT. The phenotypic correlations between pod and seed morphology related traits, and HSW and HPW ranged from 0.23 to 0.95. The highest correlations were obtained between SWI, HSW and HPW (Table 2). The plant growth habit

Table 2: Phenotypic correlations between traits in well watered (bottom left half) and water limited (upper right half). Correlations in bold are significant at 1%

	TB	PW	HW	PN	SW	HSW	HPW	SN	SHW	PWI	PL	PB	PC	SWI	SL	HI	GH	PH	PMAT	DFL
TB		0.74	0.92	0.66	0.69	0.24	0.33	0.65	0.79	0.20	0.28	0.18	-	0.16	ns	ns	-	-	ns	ns
PW	0.70		0.42	0.87	0.99	0.24	0.44	0.87	0.90	0.12	0.12	ns	-	0.16	ns	0.61	-	-	0.24	ns
HW	0.95	0.45		0.39	0.37	0.19	0.19	0.39	0.56	0.20	0.31	0.19	-	0.12	ns	-0.42	-	-	-0.15	ns
PN	0.58	0.88	0.36		0.87	-0.15	ns	0.88	0.75	-0.19	ns	ns	-	-0.17	-0.16	0.53	-	-	0.22	ns
SW	0.67	0.97	0.43	0.87		0.19	0.44	0.87	0.82	ns	ns	ns	-	0.12	ns	0.64	-	-	0.34	ns
HSW	0.23	0.21	0.20	-0.17	0.18		0.72	ns	0.36	0.66	0.51	ns	-	0.77	0.61	ns	-	-	-0.16	0.13
HPW	0.35	0.40	0.26	ns	0.38	0.80		0.20	0.44	0.54	0.35	ns	-	0.58	0.37	0.27	-	-	0.24	ns
SN	0.59	0.84	0.37	0.89	0.84	-0.16	ns		0.75	ns	ns	ns	-	ns	-0.14	0.52	-	-	ns	ns
SHW	0.69	0.91	0.47	0.77	0.82	0.27	0.41	0.73		0.31	0.29	0.13	-	0.26	0.18	0.40	-	-	ns	ns
PWI	0.19	0.16	0.17	ns	ns	0.58	0.55	ns	0.27		0.26	ns	-	0.72	0.29	ns	-	-	-0.32	0.15
PL	0.22	0.15	0.21	ns	0.12	0.43	0.41	ns	0.26	0.23		ns	-	0.25	0.67	-0.13	-	-	-0.18	ns
PB	ns	ns	ns	ns	ns	-0.13	-0.15	ns	ns	-0.20	0.20		-	ns	ns	ns	-	-	ns	ns
PC	ns	ns	0.13	ns	ns	ns	ns	ns	ns	-0.12	0.37	0.39		-	-	-	-	-	-	-
SWI	0.15	0.15	0.13	-0.20	ns	0.79	0.72	-0.12	0.26	0.69	0.31	-0.14	ns		0.46	ns	-	-	-0.18	ns
SL	0.16	ns	0.17	-0.18	ns	0.58	0.47	-0.16	0.24	0.29	0.64	0.23	0.42	0.52		ns	-	-	-0.12	ns
HI	-0.22	0.51	-0.49	0.49	0.52	ns	0.15	0.43	0.39	ns	ns	-0.12	-0.17	ns	-0.12		-	-	0.38	ns
GH	ns	-0.23	ns	-0.31	-0.20	ns	ns	-0.28	-0.24	ns	ns	ns	0.12	ns	ns	-0.23		-	-	-
PH	0.46	0.35	0.42	0.23	0.33	0.17	0.29	0.25	0.34	0.17	0.18	ns	0.14	0.14	ns	ns	0.42		-	-
PMAT	ns	0.15	ns	0.15	0.20	ns	ns	-0.14	ns	-0.25	-0.21	ns	ns	-0.17	-0.20	0.27	0.12	0.12		ns
DFL	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	-0.15	ns	

(GH) was slightly correlated with PN and SN, indicating the potential of the prostate genotypes to produce more pods and seeds. The main stem height (PH) was positively correlated with almost all traits.

QTL identification

A summary of QTLs detected in the two water regimes is provided in Table 3. At least one QTL was detected for each of the 27 traits analyzed, with a total of 95 QTLs mapped in the two environments (Figure 2).

Days to flowering

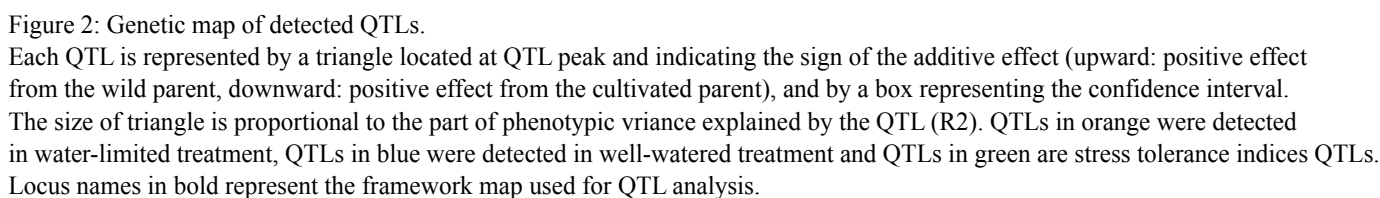
One QTL for DFL detected on LG b07, explained 9.3% of the phenotypic variance (Figure 2). Interestingly, the flowering precocity was associated with the allele of the wild amphidiploid parent.

Plant architecture

A total of 11 QTLs, explaining individually 9.8% to 26.0% of the phenotypic variance, were involved in the variation of the plant architecture (Table 3). Five QTLs located on LGs a04, a07, b04, b08 and b10 were detected for PH. Fleur 11 alleles at QTLs on LGs b04, a07, b08 and b10 tended to increase the PH, but on LG a04 the PH increase was associated with the amphidiploid allele. The five QTLs together explained 40.9% of the phenotypic variance. Interestingly, two QTLs were located on homeologous regions on LGs a04 and b04 but the parental origin of the associated allele at these two QTLs differed (Figure 2). Six QTLs located on LGs a03, a07, b04, b05, b06, and b10 were detected for GH. At the QTLs for GH, all alleles conferring an erect growth habit phenotype were from Fleur11. Overall, they explained 46.1% of the phenotypic variance. We found a colocalization of QTLs for PH and GH in three regions on LGs b04, a07 and b10.

Pod morphology

A total of 31 QTLs were detected for traits related to pod morphology when considering the two water treatments. These QTLs individually explained 8.5% to 23.9% of the phenotypic variance (Table 3). Among the 10 QTLs detected for PB, six QTLs located on LG a02, b06 and a07 were detected in the two conditions, three on LGs a08, a09 and b11 were specific to the well-watered treatment and one on LG b02 was specific to the water-limited treatment (Figure 2). Amphidiploid alleles at QTLs on LGs a02, a07, a08, a09, b02 and b11 tended to



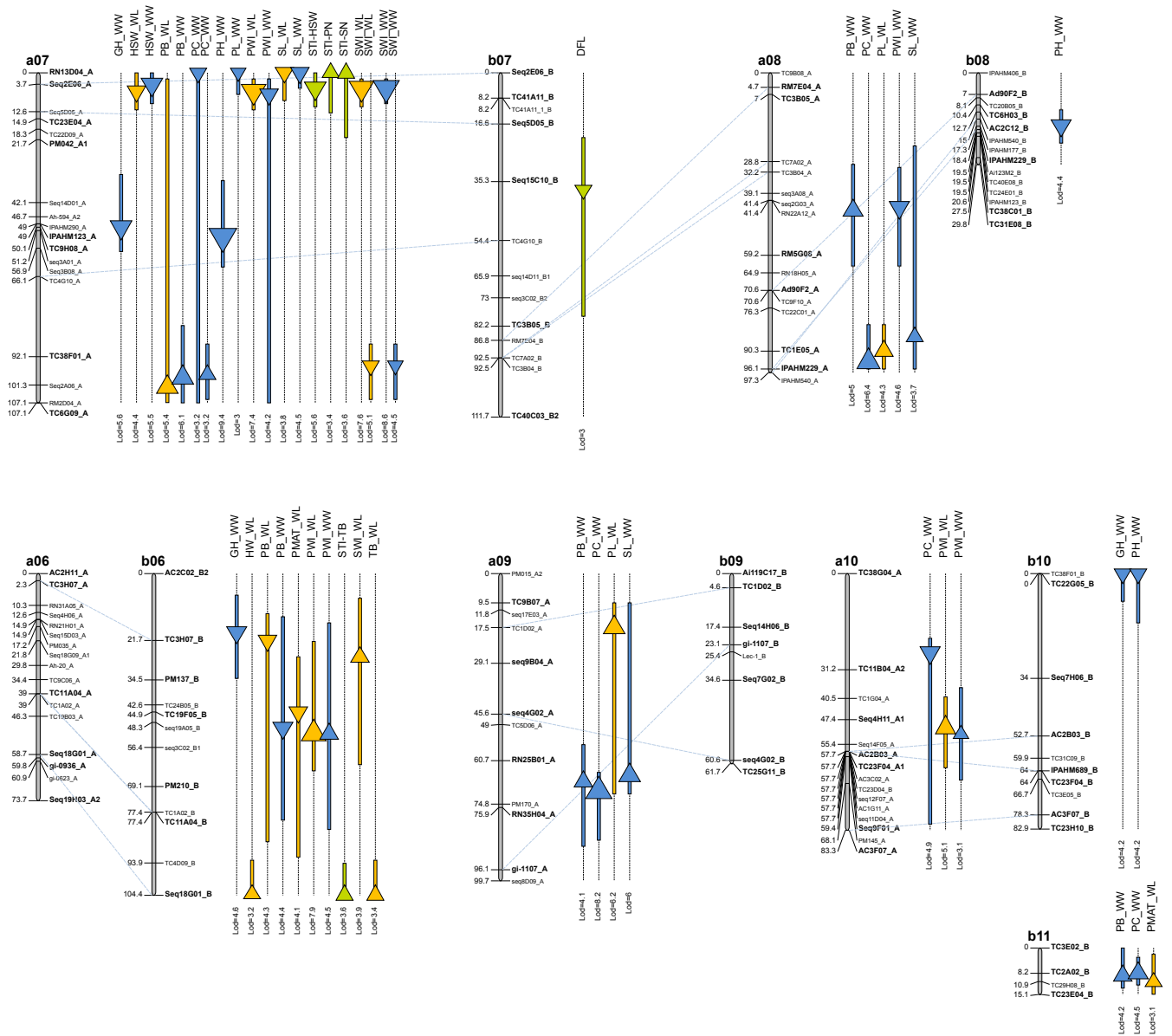


Figure 2 (Cont.)

Table 3: Characteristics of the QTLs detected for all traits in the two water treatments.

Trait category	Trait symbol	QTL Link.		Closest marker	Position	Conf. int.	R ²	Add.	Lod
		nb	group						
Flowering	DFL	1	b07	Seq15C10_B	38	21-79	9.31	-0.74	3.01
Plant architecture	GH_WW	1	a03	Seq19H03_A1	7.9	2-15	14.13	-1.19	4.29
		2	a07	TC9H08_A	50	33-58	17.30	-0.16	5.56
		3	b04	gi-0832_B	77.2	66.2-77.2	14.08	-0.71	4.61
		4	b05	Ah3TC19D09_B	0	0-7	16.17	-0.25	5.27
		5	b06	TC3H07_B	19	7-34	13.95	-0.46	4.64
		6	b10	Ah3TC22G05_B	0	0-9	9.82	-0.48	4.17
	PH_WW	1	a04	TC9E08_A1	93	44-108	9.27	2.15	2.97
		2	a07	TC9H08_A	53	35-63	26.68	-0.94	9.37
		3	b04	gi-0832_B	77.2	61.6-77.2	14.67	-2.97	4.82
		4	b08	AC2C12_B	17.1	12-22.8	13.76	-2.56	4.41
		5	b10	Ah3TC22G05_B	0	0-16	9.98	-0.66	4.2
Pod morphology	PB_WL	1	a02	seq11H01_A	15	1-65	12.63	0.28	4.33
		2	a07	TC6G09_A	102	2-107	15.42	0.42	5.38
		3	b02	TC1B02_B	5.9	0-32.9	11.82	0.30	4.04
		4	b06	TC3H07_B	21.7	13-87	12.49	-0.56	4.28
	PB_WW	1	a02	RM2H10_A	62	43-68	8.51	0.22	3.93
		2	a07	Ah3TC38F01_A	99	82-107.1	17.41	0.16	6.14
		3	a08	RM5G08_A	44.7	29.7-62.7	13.11	0.34	5.05
		4	a09	RN25B01_A	67.5	55.5-88.5	11.57	0.35	4.14
		5	b06	Ah3TC19F05_B	50	14-80	13.51	-0.08	4.35
		6	b11	TC2A02_B	9	0-13	12.67	0.45	4.16
	PC_WW	1	a02	RM2H10_A	64	43-68	23.92	1.05	8.31
		2	a07	RN13D04_A	0	0-107.1	9.98	-0.69	3.2
		3	a07	Ah3TC38F01_A	98	88-106	9.85	0.72	3.2
		4	a08	TC1E05_A	93.7	81.7-96.1	18.95	0.61	6.39
		5	a09	RN25B01_A	70.5	64.5-86.5	23.61	1.10	8.19
		6	a10	AC2B03_A	25.6	21-81.3	14.89	-0.75	4.9
	PL_WL	7	b11	TC2A02_B	8.2	3-12	13.74	0.85	4.49
		1	a08	TC1E05_A	90.3	81.7-96.1	11.98	1.29	4.34
	PL_WW	2	a09	TC9B07_A	17.5	9.5-71.5	16.51	2.09	6.16
		1	a07	RN13D04_A	0	0-7	8.03	-1.25	2.97
	PWI_WL	1	a07	seq2E06_A	6	2-12	21.10	-0.86	7.39
		2	a10	TC11B04_A2	50	40-63	15.14	0.41	5.12
	PWI_WW	3	b02	TC1B02_B	0	0-8.9	20.27	-0.78	7.07
		4	b05	PM050_B	10	1-39	14.89	-0.89	5.02
		5	b06	Ah3TC19F05_B	52	22-64	22.28	0.55	7.86
		1	a07	seq2E06_A	7	2-107.1	12.23	-0.57	4.16
		2	a08	RM5G08_A	43.7	30.7-62.7	14.36	-0.77	4.63
		3	a10	TC11B04_A2	52	37-67	8.11	0.48	3.13
		4	b02	TC1B02_B	0	0-8.9	20.13	-0.65	7.35
		5	b05	Ah3TC19D09_B	8	0-43	13.26	-0.56	4.39
		6	b06	Ah3TC19F05_B	52	16-83	12.83	0.35	4.49
		1	a07	RN13D04_A	0	0-9	11.07	-0.27	3.85
Seed morphology	SL_WL	1	a07	RN13D04_A	0	0-5	12.49	-0.13	4.47
		2	a08	TC1E05_A	85.7	23.7-96.1	10.46	0.53	3.7
		3	a09	RN25B01_A	65.5	9.5-71.5	16.27	0.83	5.97
	SWI_WL	1	a07	seq2E06_A	5	2-11	20.65	-0.46	7.56
		2	a07	Ah3TC38F01_A	95	88-106	10.12	-0.09	5.1
		3	b02	TC1B02_B	0	0-9.9	13.13	-0.37	4.59
		4	b05	seq19D06_B	35	4-46	11.04	-0.41	3.81
		5	b06	TC3H07_B	27	8-62	11.32	0.29	3.91
	SWI_WW	1	a07	seq2E06_A	5	2-10	23.69	-0.51	8.62
		2	a07	Ah3TC38F01_A	95	88-106	8.73	-0.46	4.5
		3	b02	TC1B02_B	0	0-7.9	14.22	-0.35	4.88
		4	b05	Ah3TC19D09_B	7	0-47.5	9.52	-0.32	3.18

confer a prominent beak, while on LG b06 the increase in beak prominence was associated with the allele of Fleur11. Overall, QTLs for PB explained 42.8% and 38.2% of the phenotypic variance in the well-watered and water-limited treatments, respectively. Seven QTLs located on LGs a02, a07, a08, a09, a10 and b11 were detected for PC. Amphidiploid alleles at QTLs on LGs a02, a08, a09 and b11 were associated with the constriction depth. On LG a10 the constricted pod phenotype was associated with the Fleur11 allele. The two QTLs on LG a07 were in repulsion. Overall, QTLs involved in constriction depth explained 50.2% of the phenotypic variance. Among the 11 QTLs detected for PWI, 10 QTLs located on LGs a07, a10, b02, b05 and b06 were detected in the two water treatments and one on LG a08 was specific to the well-watered treatment (Figure 2). All QTLs explained 49.2% and 58.5% of the variation of this trait in the well-watered and water-limited treatments, respectively. Fleur11 alleles at QTLs on LGs a07, a08, b02, b05 were associated with the increase in pod width while on LGs a10 and b06 the pod width increase was associated with the amphidiploid alleles. Three QTLs were detected for PL. The Fleur11 allele at the QTL on LG a07, detected in the well-watered treatment only, was associated with the increase in pod length. Amphidiploid alleles at QTLs on LG a08 and a09, detected in the water-limited treatment only, were responsible for the pod length increase.

Seed morphology

A total of 13 QTLs were detected for traits related to seed morphology when considering the two water treatments. These QTLs individually explained 8.7% to 23.0% of the phenotypic variance (Table 3). Among the nine QTLs detected for SWI, eight were detected in the two treatments, with four QTLs located on LG a07, two on LG b02 and two on LG b05. The QTL detected on LG b06 was specific to the water-limited treatment (Figure 2). Overall, these QTLs explained 38.6% and 55.5% of the phenotypic variation in the well-watered and water-limited treatments, respectively. Surprisingly, about half of the QTLs detected for SWI were distributed on LG a07, with two QTLs at a proximal position near the seq2E06_A locus and two at a distal position near the TC38F01_A locus. For QTLs on LGs a07, b02 and b05, the allele of Fleur11 conferred superior seed width in the two water treatments. The seed width increase at the QTL on LG b06 was associated with the amphidiploid allele. Four QTLs were detected for SL. Two QTLs located on LG a07 were detected in the two water treatments with the favourable alleles coming from Fleur11, while two QTLs located on LG a08 and a09 were specific to the well-watered treatment with the favourable alleles coming from the amphidiploid.

Table 3 (cont.)

Trait category	Trait symbol	QTL Link.		Closest marker	Position	Conf. int.	R ²	Add.	Lod
		nb	group						
Yield components	HI_WL	1	a02	RI2A06_A	76.1	73-76.1	18.08	-1.77	6.18
	HI_WW	1	a02	RI2A06_A	75	67-76.1	11.05	-1.07	4.18
	HPW_WL	1	b02	TC1B02_B	0	0-8.9	17.00	-12.38	6.06
		2	b05	PM050_B	7	0-52	14.96	-18.94	5.27
	HPW_WW	1	b02	TC1B02_B	0	0-10.9	20.60	-9.65	5.5
		2	b05	Ah3TC19D09_B	8	0-50	15.07	-14.90	4.63
	HSW_WL	1	a07	seq2E06_A	6	0-12	12.43	-4.86	4.42
		2	b02	TC1B02_B	0	0-10.9	14.86	-5.64	5.37
	HSW_WW	1	a07	seq2E06_A	3.7	0-10	15.74	-4.11	5.53
		2	b02	TC1B02_B	0	0-7.9	16.28	-4.21	5.72
	HW_WL	1	a02	RI2A06_A	76.1	73-76.1	13.52	2.56	4.48
		2	a05	gi-0385_A	21	0-32	17.52	4.72	5.94
		3	b06	Seq18G01_B	104	93-104.4	9.70	3.76	3.18
	HW_WW	1	a02	RI2A06_A	76	68-76.1	10.41	9.97	3.77
	PMAT_WL	1	b06	Ah3TC19F05_B	44.9	27-92	12.55	-0.01	4.11
		2	b11	TC2A02_B	11	2-15	9.62	0.02	3.1
	PMAT_WW	1	b03	PM003_B	28.1	27-48.6	9.34	0.01	3.44
	PN_WW	1	a01	TC2E05_A	4	0-26	9.27	1.96	3.03
		2	a05	gi-0385_A	17	1-34	14.20	2.00	4.88
	PW_WW	1	a01	TC2E05_A	0	1-10	11.72	2.60	3.82
	SHW_WW	1	a01	TC2E05_A	0	0-9	12.57	0.92	3.87
	SN_WW	1	a05	gi-0385_A	15	0-30	14.49	7.88	4.8
	SW_WW	1	b05	Ah3TC19E01_B	49	6-57.9	11.03	-1.91	3.58
	TB_WL	1	a05	gi-0385_A	18	0-31	16.55	7.00	5.63
		2	b06	Seq18G01_B	104.4	93-104.4	11.00	5.64	3.39
	TB_WW	1	a05	gi-0385_A	8	0-26	13.17	11.79	4.42
Stress tolerance indices	STI-HPW	1	b02	TC1B02_B	0	1-8.9	16.82	-0.16	6
		2	b05	Ah3TC19D09_B	7	0-51	13.87	-0.27	4.86
	STI-HSW	1	a07	seq2E06_A	5	0-11	15.50	-0.15	5.64
		2	b02	TC1B02_B	0	0-8.9	16.15	-0.16	5.9
	STI-HW	1	a02	RI2A06_A	76	73-76.1	16.40	0.07	5.57
		2	a05	gi-0385_A	15	0-27	17.10	0.17	5.85
	STI-PN	1	a05	gi-0385_A	22	4.6-33	19.35	0.19	6.6
		2	a07	RN13D04_A	0	0-13	10.43	0.16	3.38
	STI-PW	1	a05	gi-0385_A	18	0-33	12.34	0.16	4.14
	STI-SN	1	a07	RN13D04_A	0	0-21	10.98	0.22	3.57
	STI-SW	1	a05	gi-0385_A	19	0-34	11.51	0.13	3.78
	STI-TB	1	a05	gi-0385_A	15	0-27	20.13	0.17	7
		2	b06	Seq18G01_B	104.4	94-104.4	10.82	0.12	3.56

Yield components

A total of 26 QTLs explaining 9.2% to 20.6% of the phenotypic variance were detected for the 11 yield component traits in the two water treatments (Table 3). The number of QTLs per yield component trait ranged from one to five. Among the three QTLs significant for TB, two QTLs located on LG a05 were detected in both water treatments and one on LG b06 was specific to the water-limited treatment (Figure 2). At all these QTLs, the TB increase was associated with the amphidiploid alleles. One QTL for PW and one for SW were detected on LGs a01 and b05, respectively, in the well-watered condition. The wild allele at the QTL for PW conferred an increase in pod weight of 2.6 g per plant. For SW, the positive effect was associated with the allele of Fleur11. Four QTLs were detected for HW: two on LG a02 which were consistent across water treatments, and two on LGs a05 and b06, which were specific to the water-limited treatment. At all these QTLs, the HW increase was associated with the amphidiploid alleles. Interestingly, QTLs for HW on LGs a05 and b06 were located in the same genomic regions as those for TB, suggesting that the increase in TB could be explained by the increase in HW. QTLs for HI detected on LG a02 were detected in the two water treatments and were associated with the alleles of Fleur 11. Two QTLs located on LGs a01 and a05 for PN and one on LG a05 for SN were specific to the well-watered treatment. On LG a01, a QTL for PN colocalized with the one for PW, and on LG a05 a QTL for SN colocalized with the one for PN (Figure 2). Wild alleles at QTLs for PN and SN were responsible for an increase in pod and seed number per plant, respectively. Four QTLs were detected for HSW and four for HPW. QTLs for HSW on LGs a07 and b02 and those for HPW on LGs b02 and b05 were consistent across water treatments. QTLs for HSW and for HPW explained about 35.6% and 42.3% of phenotypic variance, respectively, in both conditions, and the positive effects were associated with the allele of Fleur11. One QTL was detected for SHW and colocalized with QTLs for PW and SN on LG a01. The amphidiploid allele at this QTL was associated with the increase in SHW. The three QTLs that conferred an increase in the percentage of pod maturity (PMAT) were detected on LG b03 in the well-watered treatment, and on LGs b06 and b11 in the water-limited treatment. The positive effects at QTLs on LGs b03 and b11 were associated with the amphidiploid alleles, while the positive effect at the QTL on LG b06 was associated with the Fleur11 allele.

QTLs related to stress tolerance indices

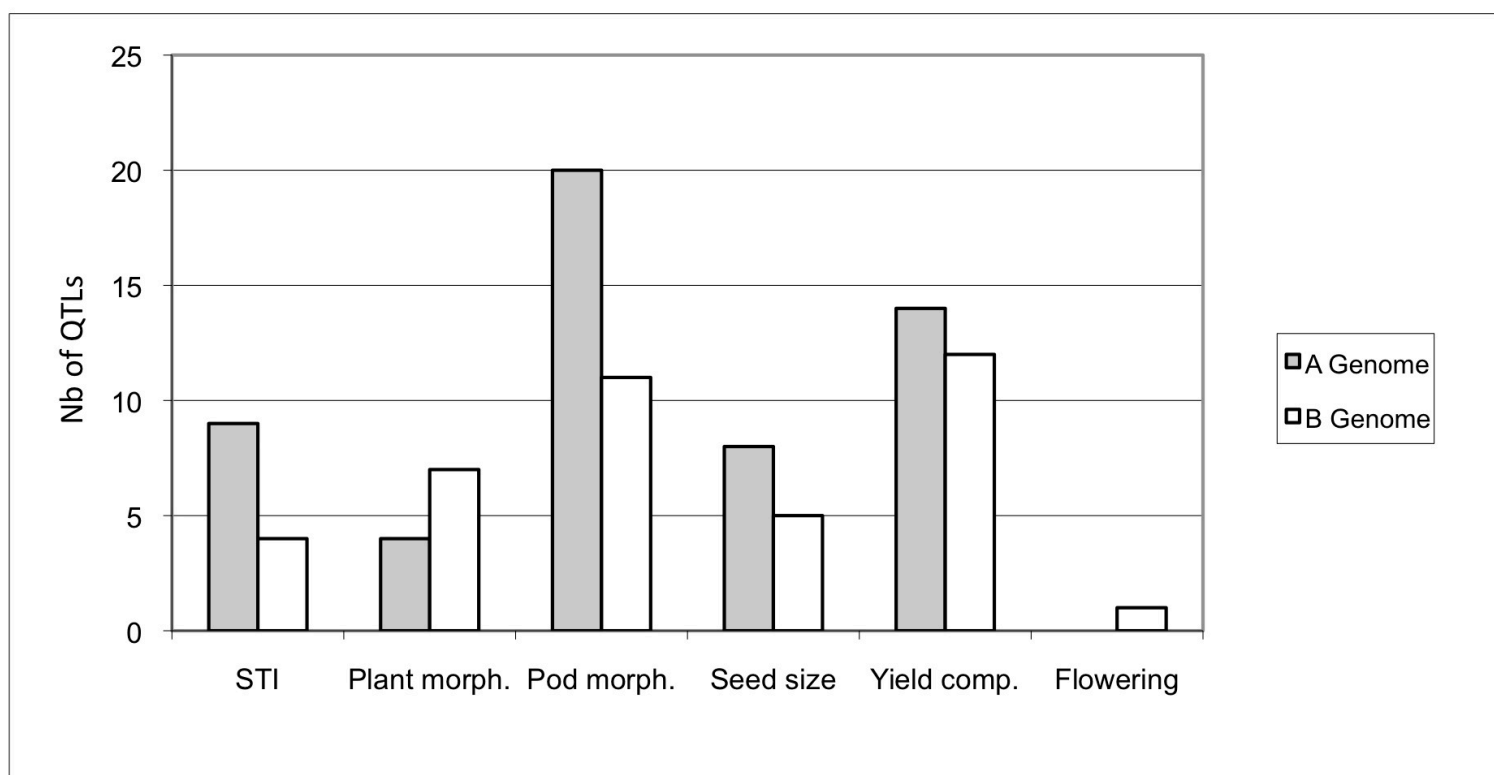


Figure 3: QTLs distribution across subgenomes. Each bar represent the number of QTLs detected in genome A (grey) and genome B (white) for each trait category (STI: stress tolerance indices, Plant morphology, Pod morphology, Seed size, Yield components, and Days to flowering).

A total of 13 QTLs were significant for the stress tolerance indices (STI): two for total biomass (STI-TB) on LGs b06 and a05, one for pod weight (STI-PW) on LG a05, one for seed weight (STI-SW) on LG a05, two for haulm weight (STI-HW) on LGs a02 and a05, two for 100 pod weight (STI-HPW) on LGs b02 and b05, two for 100 seed weight (STI-HSW) on LGs a07 and b02, two for pod number (STI-PN) on LG a05 and a07 and one for seed number (STI-SN) on LG a07. These QTLs individually explained 10.4% to 20.1% of the phenotypic variance (Table 3). In most cases, the STI related QTLs colocalized with the trait for which they were calculated. Some exceptions included QTLs for STI-PW, STI-PN and STI-SN on LG a05 and LG a07, respectively, where no QTLs for PW, PN and SN were apparently detected (Figure 2). The positive effects for QTLs for STI-HPW and STI-HSW were associated with Fleur11 alleles. For the other STI related QTLs, the positive effects were associated with the amphidiploid alleles.

Subgenomic distribution of QTLs

When taking the two water regimes into account, the number of QTLs per trait and per LG varied from 1 to 11 and from 0 to 20, respectively. We did not detect QTLs on LG a06, b01 and b09. A total of 55 QTLs were mapped on LGs belonging to the A genome, with a maximum of 20 QTLs on LG a07. For LGs belonging to the B genome, a total of 40 QTLs were detected with a maximum of 11 QTLs on LGs b02 and b06. The number of QTLs detected per trait category and per genome is shown in Figure 3. The number of QTLs globally mapped on the B subgenome was slightly lower than on the A genome (40 versus 55). We found a subgenome specific QTL contribution for some traits, including PL and SL for which QTLs were detected only on the A genome (LGs a07, a08 and a09), and DFL, PMAT and HPW for which QTLs were detected only on the B genome (b02, b03, b05, b06, b07, b11). We found a huge difference in QTL distribution both in terms of number and location on homeologous genomes. The most compelling examples were the difference in QTL number between homeologous LGs a06/b06 (0 versus 11), and between homeologous LGs a07/b07 (20 versus 1), and the difference in QTL location on homeologous LGs a02/b02 (distal versus proximal) (Figure 2). Furthermore, apart from QTLs for PB and PH that mapped to homeologous regions on LGs a02/b02 and a04/b04, all other QTLs for a given trait mapped to different homeologous LGs, thus indicating a marked inconsistency in QTL locations between homeologous LGs.

DISCUSSION

In our study, we developed an interspecific backcross population derived from a cross between Fleur11 and the amphidiploid between *A. ipaensis* and *A. duranensis*, which represent the most probable wild progenitors of allotetraploid cultivated species. This material enabled us to map several QTLs for important agronomical traits under two water regimes, to explore the reservoir of useful alleles from the wild species, analyze the subgenome contribution to the quantitative trait variation and identify chromosomal regions associated with domestication.

Peanut wild relatives represent a reservoir of useful alleles for peanut improvement

Peanut wild relatives have long been used as an important reservoir of disease resistance genes. Introgression of disease resistance genes into cultivated species using direct wild diploid x cultivated or via wild amphidiploid x cultivated crosses has been successfully used in peanut improvement (STALKER 1984; HALWARD *et al.* 1991; GARCIA *et al.* 1996). However, the use of peanut wild relatives to dissect the molecular basis of more complex traits such as yield under both normal and water limited conditions has been impeded by the lack of molecular and mapping population resources.

In our study, we used an AB-QTL mapping approach to evaluate the genetic potential of wild species to enhance important agronomical traits in cultivated peanut. We detected 95 QTLs under the two water regimes. About half of the QTL positive effects were associated with amphidiploid alleles. These QTLs, which explained a large part of the phenotypic variance, contributed positively to valuable agronomic traits such as flowering precocity (9.0%), pod weight per plant (11.7%), pod number per plant (9.0% to 14.2%), seed number per plant (14.4%), pod size (8.0% to 22.0%), seed size (11.3%), pod maturity (9.5%) and biomass production (9.0% to 17.0%). Several QTLs involved in the increase in pod and seed size, pod maturity and biomass production were specific to the water-limited treatment.

We globally observed high consistency of QTLs across water treatments. For example, among the 25 QTLs for pod and seed sizes and yield components, 17 QTLs (68.0%) were consistent in terms of genomic location across water treatments. The stability of QTLs across the water treatments may be explained by the late occurrence of severe stress during the experiment, which led to low G x E and thus QTL x E interactions. However, the terminal stress in our experiment is representative of the most common drought events in the Sahel (NDUNGURU *et al.* 1995; CLAVEL *et al.* 2004).

We used stress tolerance indices (STI) for yield components to decipher the molecular basis of yield performance under well-watered and water-stress conditions since these indices have

been considered as good criteria for identifying genotypes that combine high yield and stress tolerance potential (FERNANDEZ 1992). In addition, a positive correlation between STI for yield and yield under drought has been reported in peanut (CLAVEL *et al.* 2005; SONGSRI *et al.* 2008). Thirteen QTLs were mapped for STI related traits. Positive effects noted at nine QTLs were from the amphidiploid. Most QTLs mapped for STI related traits co-localized with the QTL of the trait for which they were calculated. This indicated a positive relationship between performances under the well-watered and water-limited treatments, as stressed above. Of special interest were STI-PN and STI-SN QTLs on LG a07, which mapped on a genomic region where no QTLs for PN and SN were mapped. Moreover, they co-localized with STI-HSW QTLs and several other QTLs involved in the increase in seed size under both well-watered and water-limited treatments. Favourable alleles at QTLs for STI-PN and STI-SN were from the amphidiploid while in this region positive alleles at all detected QTLs were from Fleur11. This suggests that this is a key region that could be involved in the trade-off between maintaining large sized seed and producing more seeds under water stress. In a cultivated allelic configuration, the production of large sized seeds could be favoured, while in a wild configuration it would be the seed number.

These results show that peanut wild relatives are valuable sources for improving important agronomical traits under both well-watered and water-limited conditions.

QTLs for the same traits are mainly found in non-homeologous regions

Globally, we detected slightly more QTLs on the A genome than on the B genome. The slight difference in QTL number across subgenomes may be related to the low level of polymorphism observed between *A. ipaensis* (wild donor of the B genome) and the B genome of the cultivated species (FONCÉKA *et al.* 2009). Interestingly, a genome specific QTL contribution was noted for PL, SL, HPW, DFL and PMAT. The most striking results were the huge differences in QTL distribution (in terms of both number and location) on homeologous LGs and the lack of QTL consistency between homeologous LGs, while good colinearity has been reported between peanut A and B genomes (BUROW *et al.* 2001; FONCÉKA *et al.* 2009; MORETZSOHN *et al.* 2009). QTLs for the same traits were mapped in 96% of cases in non-homeologous regions. Non-homeologous QTL locations could be explained by neutral and/or human driven selection of different genes in the two subgenomes that contribute to the variation in the same trait. These results could also be explained by the differential control of gene expression in subgenomes, as a consequence of interspecific hybridization and/or allotetraploidization.

Changes in gene expression, including genome specific gene silencing, unequal expression of homeologous genes, neofunctionalization or subfunctionalization of genes, have been extensively studied in various allopolyploids including wheat, cotton and brassica (DOYLE *et al.* 2008 for a review). In cotton, high variation in the expression of the A and D subgenomes have been reported (ADAMS *et al.* 2003; CHAUDHARY *et al.* 2009; FLAGEL and WENDEL 2010). In addition, RONG *et al.* (2007) reported that A and D subgenomes of the tetraploid cotton contributed QTLs for lint fiber development at largely non-homeologous locations. However, in hexaploid wheat, several authors have described the location of QTLs for the same trait on homeologous LG (SOURDILLE *et al.* 2003; QUARRIE *et al.* 2005; CHU *et al.* 2008). This suggests that variations in subgenome contributions to QTLs may depend on polyploid lineages.

Clustering of key morphological trait QTLs: footprints of domestication

The marked phenotypic differences distinguishing crops from their wild progenitors are referred to as the domestication syndrome. In legumes, mechanisms of seed dispersal (pod dehiscence), seed dormancy, growth habit, pod and seed sizes have been considered as traits contributing to the domestication syndrome (KONINANGE *et al.* 1996; GEPTS 2004; WEEDEN 2007). Peanut shares these domestication features with other grain legumes. Although pod dehiscence is absent in peanut, the long peg and isthmus observed solely in wild species have been identified as a potential mechanism for seed dispersal (PICKERSGILL 2007). In cultivated peanut pods, the isthmus is virtually nonexistent and has given way to a more or less deep pod constriction that may represent a vestige of the isthmus. Cultivated peanut also displays a more compact growth habit compared to the wild species. In general, an erect growth habit is accompanied by traits such as loss of pod constriction and of seed dormancy. This phenotype is found mainly in varieties belonging to subspecies *fastigiata*. Less-evolved growth habits are also characterized by small fruits with marked constriction and seeds demonstrating dormancy. These characters, which could be considered primitive, are mainly found in varieties belonging to the *hypogaea* subspecies (KRAPOVICKAS and GREGORY 1994).

In our study, we considered that seed and pod size (SWI, SL, PWI and PL), 100 seed and pod weights (HSW, HPW), pod constriction (PC), and plant growth habit (GH) were traits that could be mostly involved in the peanut domestication syndrome. These traits were characterized by a high heritability ranging from 0.71 to 0.90. A total of 53 QTLs were detected for these traits. More than half of them clustered in three genomic regions on LGs a07 (11 QTLs), b02 (10 QTLs), b05 (8 QTLs). All QTLs with major effects for HSW and

HPW, most of those for PWI, SWI, SL, and one for PC and GH were mapped in these three genomic regions. These QTLs individually explained 10.0% to 26.0% of the phenotypic variance and the favorable alleles were always from cultivated species. Other QTLs involved in PC and GH variations are found alone or in clusters with QTLs with lower effects for PWI, SWI, PL and SL on 10 different LGs. The high correlation between traits involved in pod size, seed size, HSW, HPW and the clustering of their associated QTLs suggests that one gene with pleiotropic effects or a limited number of linked genes are responsible for these trait variations in each region.

QTLs that greatly affected pod and seed size appeared to be clustered in three genomic regions while those affecting the plant and pod morphology seemed to be dispersed across the genome. Considering the number and distribution of QTLs for plant morphology and pod constriction on the genetic map and the wide range of peanut growth habits and constriction depths that still exist in the cultivated compartment, it is unlikely that these traits were the main focus of human selection in the incipient stages of domestication. Early stages of peanut domestication probably involved the fixation of alleles that increased pod and seed size at QTL clusters on LGs a07, b02 and b05. Preliminary modifications in plant growth habit and pod morphology could have started jointly with the increase in pod and seed sizes as QTLs for these traits were found in the same clusters. Subsequent improvements probably concerned the stacking of alleles involved in modification of the plant growth habit towards an erect type, in the modification of pod morphology, and the increase in pod and seed sizes at the other QTLs. These improvements, which could have taken place at different times and different locations, could be responsible for the morphological differences observed between peanut subspecies. Our findings on a limited number of domestication related QTLs and their clustering on specific areas of the peanut genome are in line with what has been reported for a wide range of crops, including tomato (GRANDILLO *et al.* 1999), maize (BRIGGS *et al.* 2007), rice (SWEENEY and MCCOUCH 2007), wheat (PENG *et al.* 2003), and bean (KOINANGE *et al.* 1996).

Comparative QTL analysis is a powerful tool for unraveling the genetic basis of domestication in plant families. In cereals, particularly in maize, sorghum and rice, a small number of QTLs located in syntenic regions control some domestication traits (PATERSON *et al.* 1995). In Solanaceae, DOGANLAR *et al.* (2002) reported that 40% of the loci involved in eggplant fruit weight, shape and color have putative orthologous counterparts in tomato, potato and/or pepper. However, in legumes, WEEDEN (2007) argued that although a similar number of genes have been modified during pea and common bean domestication, these

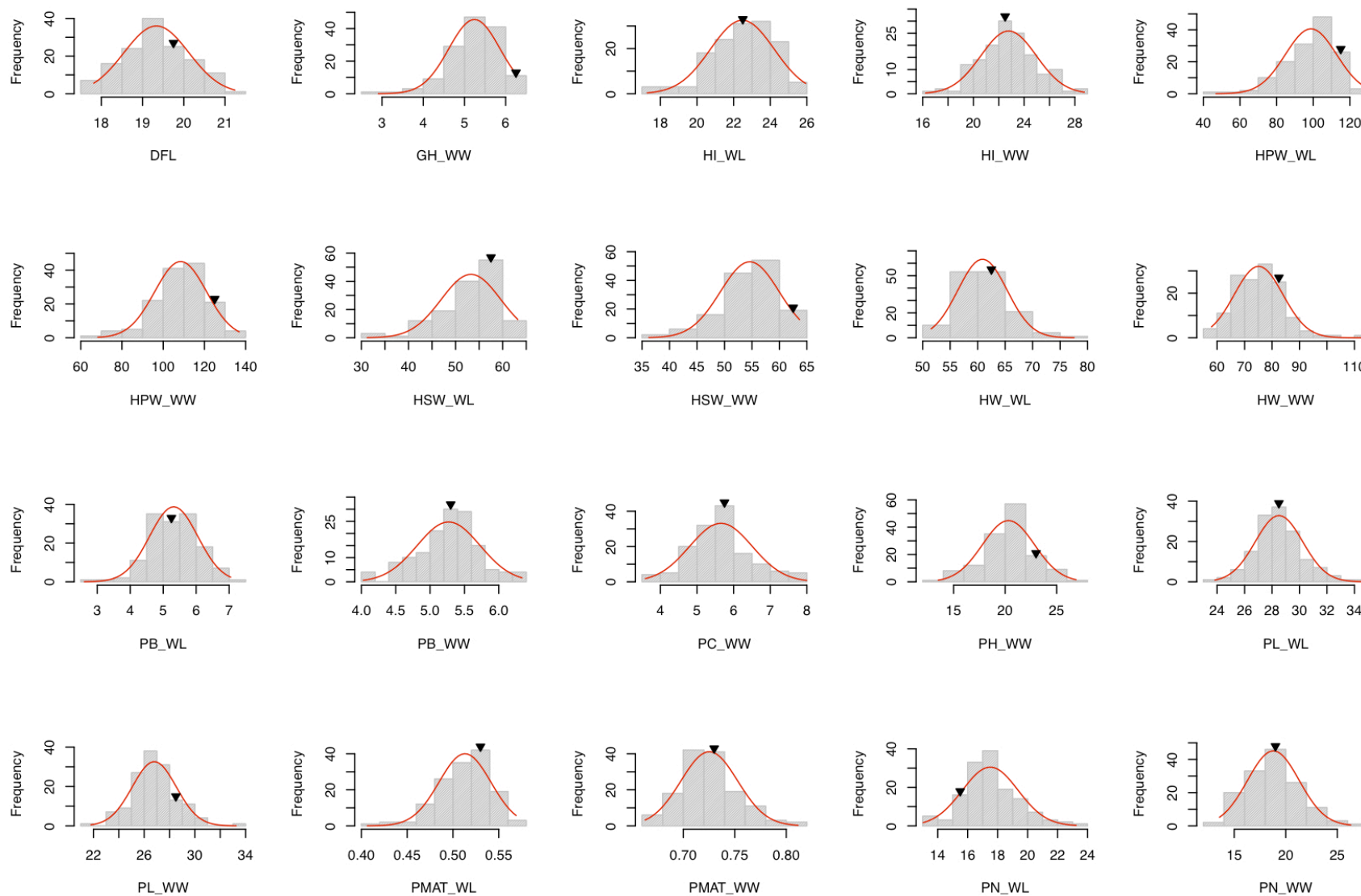


FIGURE S1. – Distribution of the traits measured in the population. The black arrow represents the value of the cultivated parent Fleur11

genes were different. Nevertheless, these authors suggested that genes responsible for seed weight, photoperiod sensitivity and seed dormancy may involve homologous or orthologous sequences. Seed weight is one of the most important traits in legume domestication (GEPTS 2004). Comparative QTL analysis across legume species showed that QTLs for seed weight were located in orthologous regions of lotus LG 2, soybean LG b1, pea LG I, and chickpea LG 8 (GONDO *et al.* 2007). A recent study on legume synteny showed that this region was also conserved between *Lotus* and the upper part of *Arachis* LG 7 (HOUGAARD *et al.* 2008), which corresponded to the seed weight QTL region in our LG a07 (data not shown). These results showed that the same orthologous regions were involved in genetic control of seed weight in peanut and several legume species and suggested that the same loci could be involved. This could be further investigated by fine mapping and cloning of these loci in different legume species.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Angelique D'Hont and David Pot for the critical review of the manuscript. This study was funded by the Cirad and the Generation Challenge Program.

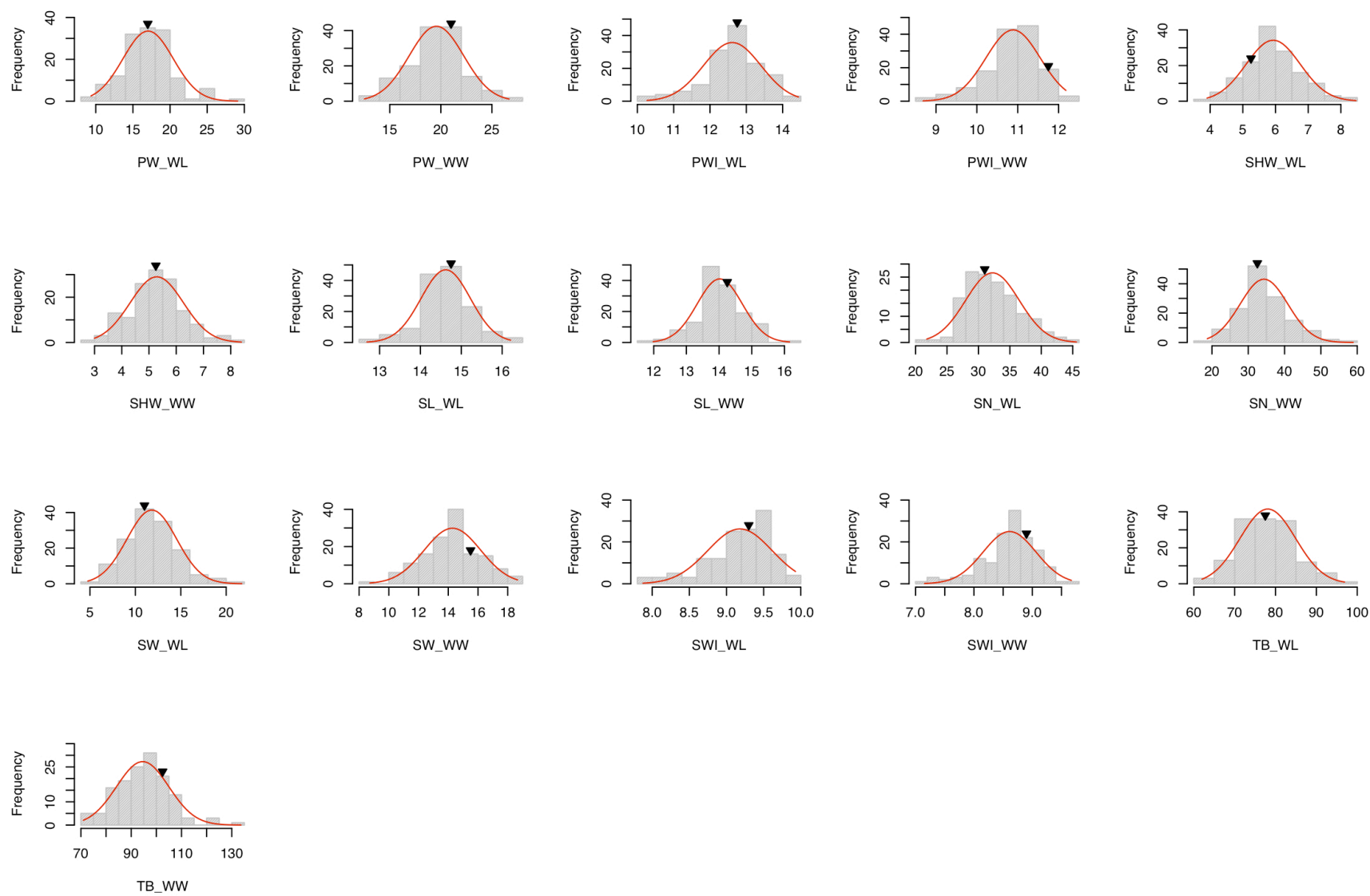


FIGURE S1. (cont.)

REFERENCES

- ADAMS, K. L., and J. F. WENDEL, 2005 Polyploidy and genome evolution in plants. *Curr Opin Plant Biol.* **8**: 135-141.
- ADAMS, K. L., R. CRONN, R. PERCIFIELD and J. F. WENDEL, 2003 Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 4649 -4654.
- BOOTE, K. J., 1982 Growth stages of peanut (*Arachis hypogaea* L.). *Peanut Science.* **9**: 35-40.
- BRIGGS, W. H., M. D. MCMULLEN, B. S. GAUT and J. DOEBLEY, 2007 Linkage mapping of domestication loci in a large maize teosinte backcross resource. *Genetics* **177**: 1915-1928.
- BROMAN, K. W., and Ś. SEN, 2009 *A Guide to QTL Mapping with R/qtl*. Springer Verlag, Berlin
- BROMAN, K. W., H. WU, Ś. SEN and G. A. CHURCHILL, 2003 R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889 -890.
- BURKE, J. M., S. TANG, S. J. KNAPP and L. H. RIESEBERG, 2002 Genetic analysis of sunflower domestication. *Genetics* **161**: 1257.
- BUROW, M. D., C. E. SIMPSON, J. L. STARR and A. H. PATERSON, 2001 Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.). Broadening the gene pool of a monophyletic polyploid species. *Genetics* **159**: 823-37.
- BUROW, M. D., C. E. SIMPSON, M. W. FARIES, J. L. STARR and A. H. PATERSON, 2009 Molecular biogeographic study of recently described B- and A-genome *Arachis* species, also providing new insights into the origins of cultivated peanut. *Genome* **52**: 107-119.
- CHAUDHARY, B., L. FLAGEL, R. M. STUPAR, J. A. UDALL, N. VERMA *et al.*, 2009 Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). *Genetics* **182**: 503-517.
- CHU, C., S. XU, T. FRIESEN and J. FARIS, 2008 Whole genome mapping in a wheat doubled haploid population using SSRs and TRAPS and the identification of QTL for agronomic traits. *Mol. Breeding* **22**: 251-266.
- CHURCHILL, G. A., AND R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963-971.
- CLAVEL, D., N. K. DRAME, H. ROY-MACAULEY, S. BRACONNIER and D. LAFFRAY, 2005 Analysis of early responses to drought associated with field drought adaptation in four

- sahelian groundnut (*Arachis hypogaea* L.) cultivars. Environ. Exper. Bot. **54**: 219-230.
- CLAVEL, D., B. SARR, E. MARONE and R. ORTIZ, 2004 Potential agronomic and physiological traits of spanish groundnut varieties (*Arachis hypogaea* L.) as selection criteria under end-of-cycle drought conditions. Agronomie **24**: 101-111.
- CUC, L., E. MACE, J. CROUCH, V. QUANG, T. LONG *et al.*, 2008 Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biol. **8**: 55.
- DOGANLAR, S., A. FRARY, M. DAUNAY, R. N. LESTER and S. D. TANKSLEY, 2002 Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. Genetics **161**: 1713-1726.
- DOYLE, J. J., L. E. FLAGEL, A. H. PATERSON, R. A. RAPP, D. E. SOLTIS *et al.*, 2008 Evolutionary genetics of genome merger and doubling in plants. Annu. Rev. Genet. **42**: 443-461.
- DWIVEDI, S. L., S. PANDE, J. N. RAO and S. N. NIGAM, 2002 Components of resistance to late leaf spot and rust among interspecific derivatives and their significance in a foliar disease resistance breeding in groundnut (*Arachis hypogaea* L.). Euphytica **125**: 81-88.
- DWIVEDI, S. L., D. BERTIOLI, J. H. CROUCH, J. F. VALLS, H. UPADHYAYA *et al.*, 2007 Peanut. pp. 115-151. in *Genome Mapping and Molecular Breeding in Plants*. edited by Kole C. Springer-Verlag, Berlin, Heidelberg.
- FAVERO, A. P., C. E. SIMPSON, J. F. M. VALLS and N. A. VELLO, 2006 Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. Crop Sci. **46**: 1546-1552.
- FAWCETT, J. A., S. MAERE and Y. VAN DE PEER, 2009 Plants with double genomes might have had a better chance to survive the cretaceous-tertiary extinction event. Proc. Natl. Acad. Sci. U.S.A. **106**: 5737-5742.
- FERGUSON, M., M. BUROW, S. SCHULZE, P. BRAMEL, A. PATERSON *et al.*, 2004 Microsatellite identification and characterization in peanut (*A. hypogaea* L.). Theor. Appl. Genet. **108**: 1064-70.
- FERNANDEZ, G., 1992 Selection criteria for assessing stress tolerance. pp. 257-269. in *International Symposium on Adaptation of Vegetables and Other Food Crops in Temperature and Water Stress*. edited by Kuo C.G., Taiïwan.
- FLAGEL, L. E., and J. F. WENDEL, 2010 Evolutionary rate variation, genomic dominance and

- duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytol.* **186**: 184-193.
- FONCÉKA, D., T. HODO-ABALO, R. RIVALLAN, I. FAYE, M. N. SALL *et al.*, 2009 Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. *BMC Plant Biol.* **9**: 103.
- GARCIA, G. M., H. T. STALKER, E. SHROEDER and G. KOCHERT, 1996 Identification of RAPD, SCAR, and RFLP markers tightly linked to nematode resistance genes introgressed from *Arachis cardenasii* into *Arachis hypogaea*. *Genome* **39**: 836-45.
- GARCIA, G. M., S. P. TALLURY, H. T. STALKER and G. KOCHERT, 2006 Molecular analysis of *Arachis* interspecific hybrids. *Theor. Appl. Genet.* **112**: 1342-1348.
- GEPTS, P., 2004 Crop domestication as a long-term selection experiment. *Plant. Breed. Rev.* **24**: 1-44.
- GIMENES, M. A., C. R. LOPES, M. L. GALGARO, J. F. M. VALLS and G. KOCHERT, 2002 RFLP analysis of genetic variation in species of section *Arachis*, genus *Arachis* (Leguminosae). *Euphytica* **123**: 421 - 429.
- GIMENES, M., A. HOSHINO, A. BARBOSA, D. PALMIERI and C. LOPES, 2007 Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*). *BMC Plant Biol.* **7**: 9.
- GONDO, T., S. SATO, K. OKUMURA, S. TABATA, R. AKASHI *et al.*, 2007 Quantitative trait locus analysis of multiple agronomic traits in the model legume *Lotus japonicus*. *Genome* **50**: 627-637.
- GRANDILLO, S., H. M. KU and S. D. TANKSLEY, 1999 Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theor. Appl. Genet.* **99**: 978-987.
- GUR, A., and D. ZAMIR, 2004 Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol.* **2**: e245.
- HAJJAR, and HODGKIN, 2007 The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* **156**: 1-13.
- HALWARD, T. M., J. C. WYNNE and H. T. STALKER, 1991 Recurrent selection progress in a population derived from an interspecific peanut cross. *Euphytica* **52**: 79-84.
- HAMMER, K., 1984 Das domestikationssyndrom. *Die Kulturpflanze* **32**: 11-34.
- HARLAN, J. R., 1992 Origins and processes of domestication. pp. 159–175. in *Grass evolution and domestication*. edited by G.P. Chapman, Cambridge University Press, Cambridge.
- HE, G., R. MENG, M. NEWMAN, G. GAO, R. PITTMAN *et al.*, 2003 Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). *BMC Plant Biol.* **3**: 3.

- HE, G., and C. S. PRAKASH, 1997 Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). *Euphytica* **97**: 143 - 149.
- HONG, Y., X. CHEN, X. LIANG, H. LIU, G. ZHOU *et al.*, 2010 A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. *BMC Plant Biol.* **10**: 17.
- HOPKINS, M. S., A. M. CASA, T. WANG, S. E. MITCHELL, R. E. DEAN *et al.*, 1999 Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Sci.* **39**: 1243-1247.
- HOUGAARD, B. K., L. H. MADSEN, N. SANDAL, M. DE CARVALHO MORETZSOHN, J. FREDSLUND *et al.*, 2008 Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics* **179**: 2299-2312.
- IBPGR and ICRISAT, 1992. *Descriptors for groundnut*. edited by Int. Board of Plant Genetic Resources and Int. Crops Res. Inst. For the Semi-Arid Tropics, Rome, Italy and Patancheru, A.P., India.
- KOCHERT, G., H. T. STALKER, M. GIMENES, M. L. GALGARO, C. R. LOPES *et al.*, 1996 RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *Am. J. Bot.* **83**: 1282-1291.
- KOINANGE, E. M. K., S. P. SINGH and P. L. GEPTS, 1996 Genetic control of the domestication syndrome in common bean. *Crop Sci.* **36**: 1037–1045.
- KRAPOVICKAS, A., and W. GREGORY, 1994 Taxonomía del género *Arachis* (Leguminosae). *Bonplandia* **8**: 1-186.
- LEAL-BERTIOLI, S., A. C. JOSE, D. ALVES-FREITAS, M. MORETZSOHN, P. GUIMARAES *et al.*, 2009 Identification of candidate genome regions controlling disease resistance in *Arachis*. *BMC Plant Biol.* **9**: 112.
- LIANG, X., X. CHEN, Y. HONG, H. LIU, G. ZHOU *et al.*, 2009 Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biol.* **9**: 35.
- LU, J., and B. PICKERSGILL, 1993 Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L.-Leguminosae). *Theor Appl Genet.* **85**: 550-560.
- MILLA, S. R., T. G. ISLEIB and H. T. STALKER, 2005 Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. *Genome* **48**: 1-11.
- MORETZSOHN, M., M. HOPKINS, S. MITCHELL, S. KRESOVICH, J. VALLS *et al.*, 2004 Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis

- of hypervariable regions of the genome. *BMC Plant Biol.* **4**: 11.
- MORETZSOHN, M., L. LEOI, K. PROITE, P. GUIMARÃES, S. LEAL-BERTIOLI *et al.*, 2005 A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theor. Appl. Genet.* **111**: 1060-1071.
- MORETZSOHN, M., A. BARBOSA, D. ALVES-FREITAS, C. TEIXEIRA, S. LEAL-BERTIOLI *et al.*, 2009 A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. *BMC Plant Biol.* **9**: 40.
- NDUNGURU, B. J., B. R. NTARE, J. H. WILLIAMS and D. C. GREENBERG, 1995 Assessment of groundnut cultivars for end-of-season drought tolerance in a sahelian environment. *J. Agr. Sci.* **125**: 79–85.
- PANDE, S., and J. N. RAO, 2001 Resistance of wild *Arachis* species to late leaf spot and rust in greenhouse trials. *Plant Dis.* **85**: 851-855.
- PATERSON, A. H., K. F. SCHERTZ, Y. R. LIN, S. C. LIU and Y. L. CHANG, 1995 The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) pers. *Proc. Natl. Acad. Sci. U. S. A.* **92**: 6127-31.
- PATERSON, A. H., 2002 What has QTL mapping taught us about plant domestication? *New Phytol.* **154**: 591-608.
- PATERSON, A. H., 2005 Polyploidy, evolutionary opportunity, and crop adaptation. *Genetica* **123**: 191-196.
- PENG, J., Y. RONIN, T. FAHIMA, M. S. RÖDER, Y. LI *et al.*, 2003 Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc. Natl. Acad. Sci. U.S.A.* **100**: 2489 -2494.
- PICKERSGILL, B., 2007 Domestication of plants in the Americas: insights from mendelian and molecular genetics. *Ann. Bot.* **100**: 925-940.
- PROITE, K., S. C. M. LEAL-BERTIOLI, D. J. BERTIOLI, M. C. MORETZSOHN, F. R. D. SILVA *et al.*, 2007 ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC Plant Biol.* **7**: 7.
- PURUGGANAN, M. D., and D. Q. FULLER, 2009 The nature of selection during plant domestication. *Nature* **457**: 843-848.
- QUARRIE, S. A., A. STEED, C. CALESTANI, A. SEMIKHODSKII, C. LEBRETON *et al.*, 2005 A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* **110**: 865-880.

- R DEVELOPMENT CORE TEAM, 2010. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.
- RAINA, S. N., V. RANI, T. KOJIMA, Y. OGIHARA, K. P. SINGH *et al.*, 2001 RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome* **44**: 763-72.
- REDDY T. Y., V. R. REDDY and V. ANBUMOZHI, 2003 Physiological responses of groundnut (*Arachis hypogaea* L.) to drought stress and its amelioration: a critical review. *Plant Growth Regul.* **41**: 75-88.
- REDDY, A. S., L. J. REDDY, N. MALLIKARJUNA, M. D. ABDURAHMAN, Y. V. REDDY *et al.*, 2000 Identification of resistance to peanut bud necrosis virus (PBNV) in wild *Arachis* germplasm. *Ann. Appl. Biol.* **137**: 135-139.
- ROBLEDO, G., G. LAVIA and G. SEIJO, 2009 Species relations among wild *Arachis* species with the a genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. *Theor. Appl. Genet.* **118**: 1295-1307.
- ROBLEDO, G., and G. SEIJO, 2010 Species relationships among the wild B genome of *Arachis* species (section *Arachis*) based on fish mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. *Theor. Appl. Genet.* **121**: 1033-1046.
- RONG, J., F. A. FELTUS, V. N. WAGHMARE, G. J. PIERCE, P. W. CHEE *et al.*, 2007 Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. *Genetics* **176**: 2577-2588.
- SADRAS, V. O., and S. P. MILROY, 1996 Soil-water thresholds for the responses of leaf expansion and gas exchange: a review. *Field Crop. Res.* **47**: 253-266.
- SARR, B., J. LECOEUR and P. CLOUVEL, 2004 Irrigation scheduling of confectionery groundnut (*Arachis hypogaea* L.) in Senegal using a simple water balance model. *Agr. Water Manage.* **67**: 201-220.
- SEIJO, G., G. I. LAVIA, A. FERNANDEZ, A. KRAPOVICKAS, D. A. DUCASSE *et al.*, 2007 Genomic relationships between the cultivated peanut (*Arachis hypogaea*, Leguminosae) and its close relatives revealed by double gish. *Am. J. Bot.* **94**: 1963-1971.
- SERVIN, B., C. DILLMANN, G. DECOUX and F. HOSPITAL, 2002 MDM: a program to compute fully informative genotype frequencies in complex breeding schemes. *J. Hered.* **93**: 227-8.

- SHARMA, S., M. ANSARI, K. VARAPRASAD, A. SINGH and L. REDDY, 1999 Resistance to *Meloidogyne javanica* in wild *Arachis* species*. Genet. Resour. Crop Ev. **46**: 557-568.
- SIMPSON, C., and J. STARR, 2001 Registration of 'COAN' peanut. Crop Sci. **41**: 918.
- SIMPSON, C., J. STARR, G. CHURCH, M. BUROW and A. PATERSON, 2003 Registration of 'NEMATAM' peanut. Crop Sci. **43**: 1561.
- SINGH, K., A. SINGH, S. RAINA, A. SINGH and Y. OGIHARA, 2002 Ribosomal DNA repeat unit polymorphism and heritability in peanut (*Arachis hypogaea* L.) accessions and related wild species. Euphytica **123**: 211-220.
- SMITH, B. D., 2001 Documenting plant domestication: the consilience of biological and archaeological approaches. Proc. Natl. Acad. Sci. U.S.A. **98**: 1324-1326.
- SOLTIS, D. E., V. A. ALBERT, J. LEEBENS-MACK, C. D. BELL, A. H. PATERSON *et al.*, 2009 Polyploidy and Angiosperm diversification. Am. J. Bot. **96**: 336-348.
- SOLTIS, P. S., and D. E. SOLTIS, 2000 The role of genetic and genomic attributes in the success of polyploids. Proc. Natl. Acad. Sci. U.S.A. **97**: 7051-7057.
- SONGSRI, P., S. JOGLOY, N. VORASOOT, C. AKKASAENG, A. PATANOTHAI *et al.*, 2008 Root distribution of drought-resistant peanut genotypes in response to drought. J. Agron. Crop Sci. **194**: 92-103.
- SOURDILLE, P., T. CADALEN, H. GUYOMARC'H, J. W. SNAPE, M. R. PERRETANT *et al.*, 2003 An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. Theor. Appl. Genet. **106**: 530-538.
- STALKER, H. T., 1984 Utilizing *Arachis cardenasii* as a source of *Cercospora* leafspot resistance for peanut improvement. Euphytica **33**: 529-538.
- SWAMY, B. P. M., and N. SARLA, 2008 Yield-enhancing quantitative trait loci (QTLs) from wild species. Biotechnol. Adv. **26**: 106-120.
- SWEENEY, M., and S. MCCOUCH, 2007 The complex history of the domestication of rice. Ann Bot. **100**: 951-957.
- TALLURY, S. P., K. W. HILU, S. R. MILLA, S. A. FRIEND, M. ALSAGHIR *et al.*, 2005 Genomic affinities in *Arachis* section *Arachis* (Fabaceae): molecular and cytogenetic evidence. Theor. Appl. Genet. **111**: 1229 - 1237.
- TANKSLEY, S. D., S. GRANDILLO, T. M. FULTON, D. ZAMIR, Y. ESHED *et al.*, 1996 Advanced Backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. Theor. Appl. Genet. **92**: 213-224.
- TANKSLEY, S. D., and S. R. MCCOUCH, 1997 Seed banks and molecular maps: unlocking genetic potential from the wild. Science **277**: 1063-1066.

- UDALL, J. A., and J. F. WENDEL, 2006 Polyploidy and crop improvement. *Crop Sci.* **46**: S-3-14.
- VARSHNEY, R., D. BERTIOLI, M. MORETZSOHN, V. VADEZ, L. KRISHNAMURTHY *et al.*, 2008 The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* **118**: 729-739.
- WEEDEN, N. F., 2007 Genetic changes accompanying the domestication of *Pisum sativum*: is there a common genetic basis to the 'domestication syndrome' for legumes? *Ann Bot.* **100**: 1017-1025.
- WENDEL, J. F., 2000 Genome evolution in polyploids. *Plant Mol. Biol.* **42**: 225-249.
- ZEDER, M. A., E. EMSHWILLER, B. D. SMITH and D. G. BRADLEY, 2006 Documenting domestication: the intersection of genetics and archaeology. *Trends Genet.* **22**: 139-155.

Chapitre IV CONSTRUCTION D'UNE COLLECTION DE LIGNEES D'INTROGRESSION A PARTIR D'UN CROISEMENT INTERSPECIFIQUE CHEZ L'ARACHIDE.

I. BREF RAPPEL DE LA PROBLEMATIQUE

L'arachide est une culture importante de la zone sahélienne. Elle y est cultivée comme culture de rente, pour la production d'huile, et également pour l'utilisation fourragère des résidus de récolte.

L'arachide cultivée *Arachis hypogaea* est une des 80 espèces connues du genre *Arachis*. Sa structure allotétraploïde est attribuée à un événement relativement récent d'hybridation entre deux espèces diploïdes sauvages, suivi d'un événement de duplication spontanée du nombre de chromosomes (Kochert et al., 1996; Seijo et al., 2004).

L'espèce cultivée est isolée, d'un point de vue reproductif, des espèces diploïdes sauvages. Il en résulte une base génétique relativement étroite et une faible variabilité à la disposition des sélectionneurs pour aborder les principales contraintes de la culture, parmi lesquelles, la sensibilité à plusieurs ravageurs et maladies ou le stress hydrique. Le faible niveau de polymorphisme observé au sein de l'espèce cultivée a freiné le développement et l'utilisation d'outils moléculaires. L'élargissement de la base génétique de l'espèce cultivée dans les programmes de sélection par le recours aux espèces sauvages a été jusqu'à présent limité par cette barrière reproductive bien que des sources de résistance ainsi que des caractères d'adaptation aient été identifiés dans diverses espèces sauvages (Rao and Bramel, 2003).

L'utilisation de variétés amphidiploïdes synthétiques, obtenues par croisement de deux espèces sauvages diploïdes et par doublement du stock chromosomique permet de contourner cette difficulté. En effet, les variétés obtenues, composées de deux génomes sauvages, produisent des hybrides fertiles lorsqu'elles sont croisées avec l'arachide cultivée. Elles permettent de disposer de populations expérimentales qui constituent un matériel de choix pour l'étude des caractères d'intérêt agronomique.

Les lignées de substitution de segments chromosomiques (CSSL), sont particulièrement adaptées à l'étude des caractères complexes dans le cadre d'introgressions en provenance de matériel exotique (Gur and Zamir, 2004; Grandillo et al., 2007). Elles constituent un matériel fixé qui peut être multiplié, distribué et caractérisé dans plusieurs environnements. La puissance de détection de QTL y est optimisée par rapport aux approches classiques de

Tableau IV.1 : Description des variétés cultivées utilisées comme parent récurrent pour la production de populations BC₁F₁

Variétés	Sous-espèce	Type botanique	Type commercial	Cycle semis-maturité (j)	Port de la plante	Poids de 100 graines (g)	Rendement potentiel (t/ha)	Autres caractères
55-437	<i>fastigiata</i>	<i>vulgaris</i>	spanish	90	érigé	35-38	2	sensible à la cercosporiose et aux pucerons, tolérante à la sécheresse, faible dormance des graines
Fleur11	<i>fastigiata</i>	<i>vulgaris</i>	spanish	90-95	érigé	50-55	2	sensible à la cercosporiose, à la rouille et aux pucerons, tolérante à la sécheresse, pas de dormance des graines
GC8-35	<i>fastigiata</i>	<i>vulgaris</i>	spanish	80	érigé	30-35	1.5	variété très précoce, production de fanes moyenne, pas de dormance des graines
73-30	<i>fastigiata</i>	<i>vulgaris</i>	spanish	90-95	érigé	38-40	2	sensible à la rosette et aux pucerons, tolérante à la sécheresse, forte dormance des graines

Tableau IV.2 : Description des espèces sauvages diploïdes à l'origine des amphidiploïdes utilisés comme parent

Nom	Accession	Génome	Cycle	Port de la plante	Hauteur tige principale (cm)	Longueur axe secondaire (cm)	Poids de 100 graines (g)
<i>A. ipaensis</i>	KG30076	B	annuel	rampant	45	50	28.8
<i>A. duranensis</i>	V14167	A	annuel	rampant	20-30	140	13.12
<i>A. diogoi</i>	GKP10602	A	perenne	rampant	-	30-60	12.89
<i>A. cardenasii</i>	GKP10017	A	perenne	rampant	10-90	60-120	-
<i>A. batizocoi</i>	K9484	B	annuel	rampant	30-100	400	22

détection dans des populations en ségrégation car l'effet des allèles en provenance du donneur est évalué dans le fonds génétique constant d'une variété cultivée. Enfin, ces lignées sont très utiles comme point de départ de travaux de création variétale par pyramidage de QTL dans un même fonds génétique (Gur and Zamir, 2004) et de clonage de QTL (Yano, 2001), puisque des lignées isogéniques peuvent en être facilement dérivées.

Ce travail de développement de lignées d'introgression fait suite aux deux premiers chapitres sur la construction de la carte génétique et la cartographie de QTL dans les populations de backcross avancé. Il a pour objectif de créer un matériel génétique plus générique qui pourra être utilisé tant pour des applications de création variétale que pour celles de cartographie fine de nouveaux allèles provenant du compartiment sauvage apparenté à l'arachide cultivée. Ce chapitre présente les stratégies et les différentes étapes mises en place pour le développement d'une population de lignées d'introgression à partir d'un croisement interspécifique chez l'arachide.

II. MATERIEL ET METHODES

Les parents récurrents

Quatre variétés cultivées appartenant au type commercial spanish, variété botanique *vulgaris*, sous-espèce *fastigiata* ont été utilisées comme parent récurrent pour produire 5 populations BC₁F₁. Il s'agit des variétés 55-437, Fleur11, GC8-35 et 73-30. Les caractéristiques de ces variétés figurent dans le Tableau IV.1.

Les parents donneurs sauvages

Deux tétraploïdes synthétiques ou amphidiploïdes issus du croisement entre des espèces diploïdes sauvages ont été utilisés comme parent donneur. Le premier amphidiploïde, AiAd, résulte du croisement entre *A. ipaensis* KG30076 (génom B) et *A. duranensis* V14167 (génom A) (Favero et al., 2006). Le second amphidiploïde, TxAG6, résulte d'un croisement plus complexe impliquant *A. diogoi* et *A. cardenasii* (génom A) et *A. batizocoi* (génom B) (Simpson et al. 1993). Une description des espèces diploïdes sauvages à l'origine des amphidiploïdes figurent dans le Tableau IV.2.

Développement des populations BC₁F₁



Les étapes du croisement de l'arachide.

- a. Castration de la fleur
- b. Prélèvement du sac pollinique sur une fleur de la plante mâle
- c. Dépôt du sac pollinique sur une fleur de la plante femelle
- d. Fleur croisée

Le développement des populations BC₁F₁ a été réalisé en serre au CERAAS entre mars 2006 et juillet 2007. Les croisements ont été effectués entre l'amphidiploïde AiAd et les variétés 55-437, Fleur11, GC8-35 et 73-30, et entre l'amphidiploïde TxAG6 et la variété GC8-35.

Les précédentes expériences de multiplication avaient montré que les amphidiploïdes fleurissaient environ un mois plus tard que les variétés cultivées. Dans l'objectif de synchroniser les périodes de floraison, les amphidiploïdes, utilisés comme parent mâles, ont été semés un mois avant les variétés cultivées. Les fleurs d'une dizaine de plantes de chaque amphidiploïdes ont été utilisées pour féconder dix plantes par variété utilisée comme parent récurrent de chaque population. Les croisements ont été réalisés suivant le protocole standard d'hybridation utilisé par le programme de sélection de l'arachide de l'Institut Sénégalais de Recherche Agricole (ISRA). Nous en rappelons brièvement les principales étapes. Une castration manuelle des fleurs non encore écloses des parents utilisés comme femelle a été réalisée la veille de chaque croisement. Le lendemain, à l'aube, les sacs polliniques des fleurs des plantes utilisées comme parent mâle ont été placés sur les styles des fleurs femelles préalablement émasculées. Les croisements ont été réalisés entre 6 et 10 heures du matin. Toutes les fleurs qui n'ont pu être fécondées ont été supprimées manuellement de manière à réduire les risques d'auto ou d'allopollinisation non contrôlées.

Les graines F₁ putatives ont été récoltées en bulk sur les dix plantes puis 10 à 15 graines de chaque croisement ont été ressemées. En dépit des précautions qui sont prises pour réaliser les croisements, il existe toujours un risque élevé d'autofécondation chez l'arachide. Par conséquent, des caractères morphologiques tels que la vigueur et le port des plantes ont été utilisés pour différencier les plantes réellement issues d'une hybridation de celles issues d'un événement non-contrôlé d'autofécondation. Pour chaque croisement, les plantes identifiées comme de vraies F₁ ont été utilisées comme parent mâles en rétrocroisement avec une dizaine de plantes de chaque parent récurrent. Les graines BC₁F₁ putatives ont été récoltées en bulk sur les dix plantes.

Criblage du polymorphisme de marqueurs microsatellites sur les parents des différentes populations

Le criblage du polymorphisme des marqueurs microsatellites s'est effectué au CIRAD sur un panel comprenant les deux espèces diploïdes, *A. ipaensis* et *A. duranensis*, les amphidiploïdes AiAd et TxAG6 et les parents cultivés 55-437, Fleur11, GC8-35 et 73-30. Ce criblage s'est effectué en trois phases liées à la disponibilité des marqueurs dans le domaine public. Les protocoles d'extraction d'ADN et d'amplification ont été décrits dans le chapitre sur la

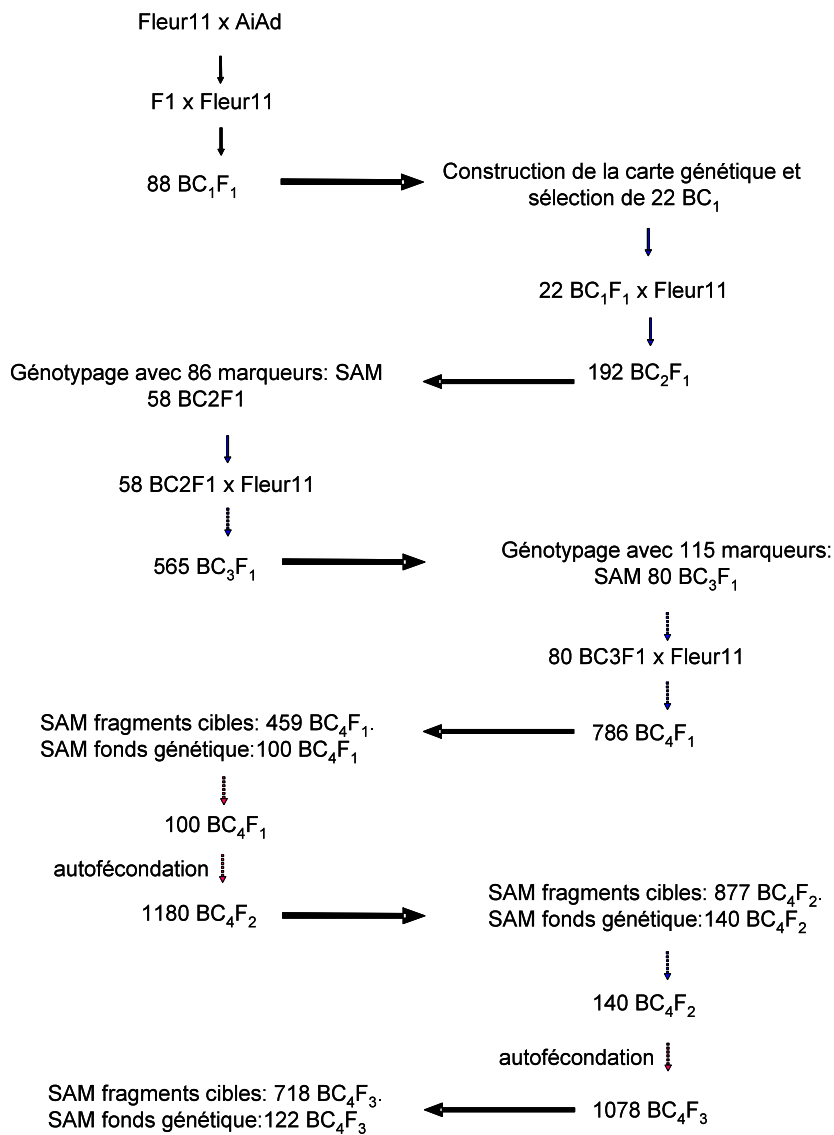


Figure IV.1 : Schéma de sélection mis en œuvre pour le développement des ILs. SAM : sélection assistée par marqueurs.

cartographie génétique. En octobre 2007, un premier jeu de 234 marqueurs a été testé sur le panel. Le nombre de marqueurs polymorphes étant insuffisant pour répondre aux objectifs de cartographie génétique et de construction des populations de lignées d'introgession, un second jeu de 172 marqueurs puis un troisième jeu de 181 marqueurs ont été utilisés en mars 2008 et octobre 2008. Au total, 587 marqueurs microsatellites ont été évalués.

Contrôle des populations BC_1F_1 et choix de la population d'étude

Dans l'objectif d'estimer le taux de réussite des croisements, un contrôle de 12 individus pris au hasard dans chacune des 5 populations BC_1F_1 a été effectué à l'aide de 12 marqueurs microsatellites. Les critères pour le choix de la population d'étude ont porté sur le taux de réussite des croisements, le nombre d'individus disponibles dans chaque population BC_1F_1 , le niveau de polymorphisme entre les parents et les caractéristiques physiologiques et agronomiques des parents récurrents.

Développement de la collection de lignées d'introgession

Les aspects liés au développement de la carte génétique et à l'étude des introgressions de fragments d'ADN en provenance des espèces sauvages dans les populations BC_1F_1 et BC_2F_1 ont été traités en partie dans les chapitres précédents. Nous abordons dans ce paragraphe quelques points qui n'ont pas été pris en compte dans les précédents chapitres et qui ont été déterminants dans le choix de la stratégie que nous avons adoptée pour produire les différentes générations de backcross. De plus, en raison du nombre limité de marqueurs microsatellites disponibles au début de ces travaux, dans les trois premières générations de backcross nous avons conduit en parallèle les activités de construction des populations et celle de saturation de la carte génétique. Enfin, les variétés choisies étant de cycle court, le temps qui sépare le semis à l'apparition des premières fleurs est de 21 jours ce qui laisse un temps relativement limité pour le choix des individus à avancer en génération suivante. Le schéma général de sélection qui a été mis en œuvre pour développer les lignées d'introgession est indiqué dans la Figure IV.1.

Choix des individus BC_1F_1 pour développer la génération BC_2F_1 .

Fin janvier 2008, 113 graines BC_1F_1 issus du croisement Fleur11 x AiAd et 30 graines du parent récurrent (Fleur11) ont été semées en serre au CERAAS. Quinze jours après semis, de jeunes feuilles ont été prélevées sur chaque BC_1F_1 et chaque individu Fleur11, et les ADN ont été extraits au CERAAS. Un premier criblage des 113 individus BC_1F_1 à l'aide de 12

marqueurs microsatellites a été effectué au CIRAD. Sur la base de ce criblage 88 BC₁F₁ conformes ont été identifiés. Ce criblage a aussi révélé que 42 BC₁F₁ présentaient des allèles non-parentaux à au moins 2 loci. Ces individus n'ont pas été retenus pour le développement de la génération BC₂F₁. Le contrôle des parents récurrents avec les mêmes douze marqueurs microsatellites a permis d'attribuer l'origine des allèles non-parentaux à de l'hétérogénéité résiduelle au niveau du lot de semences de Fleur11 utilisé. Les 46 BC₁F₁ ne présentant aucun allèle non-parental ont été utilisés comme parents mâles pour être croisés avec des individus Fleur11 homogènes choisis parmi les 30 plantes contrôlées. Pour la suite des activités de développement des populations, nous n'avons utilisé comme parent récurrent que des individus Fleur11 contrôlés.

En parallèle, les 88 BC₁F₁ ont été génotypés avec 132 marqueurs polymorphes pour produire une première carte génétique comprenant 20 groupes de liaison et une taille totale de 1487.8 cM. Pour choisir parmi les 46 BC₁F₁ un nombre minimal d'individus possédant des fragments chromosomiques du génome donneur qui permettent une couverture totale du génome du parent récurrent, une carte de travail un peu moins dense comprenant 124 marqueurs a été définie. Le logiciel CSSL-Finder (Lorieux, 2007 <http://mapdisto.free.fr/CSSLFinder/>) a été utilisé avec les critères de choix suivants : une taille de fragment donneur égale à 30 cM, un chevauchement des fragments adjacents contenus dans différentes lignées consécutives et le meilleur retour possible au fonds génétique du parent récurrent. Nous avons choisi de construire un jeu de données dans lequel chaque individu BC₁F₁ était représenté plusieurs fois dans l'objectif de transférer les fragments cibles d'un même individu BC₁F₁ dans plusieurs individus BC₂F₁. A l'aide de cette stratégie, un jeu de 22 individus BC₁F₁ a été retenu et 364 graines BC₂F₁ ont été récoltées sur ces individus.

Choix des individus BC₂F₁ pour développer la population BC₃F₁.

Trois cents graines BC₂F₁ représentant 22 familles BC₁F₁ (10 à 15 graines par plante) ont été semées en serre en juin 2008. Un taux de germination de 64% a été enregistré en raison du manque de maturité et de la dormance de certaines graines. Au total, les ADN de 192 individus ont été extraits au CERAAS. Entre mars et juin 2008 un criblage de 172 marqueurs microsatellites avait permis de placer de nouveaux marqueurs sur la carte génétique la faisant passer ainsi à 206 marqueurs et à 21 groupes de liaison d'une taille totale de 1515.5 cM. A partir de cette nouvelle carte génétique, un jeu de 86 marqueurs régulièrement répartis, avec une densité d'environ 1 marqueur tous les 17 cM, a été utilisé pour génotyper les 192 individus BC₂F₁. La même stratégie que celle décrite précédemment a été adoptée pour

choisir 58 individus BC_2F_1 à avancer en génération BC_3F_1 . Pour le développement de la population BC_3F_1 , les 58 individus BC_2F_1 ont été utilisés comme parents femelles pour être croisés avec le parent récurrent et 1207 graines BC_3F_1 ont été récoltées fin septembre 2008.

Choix des individus BC_3F_1 pour développer la population BC_4F_1 .

En raison d'une forte pression parasitaire dans les serres durant la création de la population BC_3F_1 , le développement de la population BC_4F_1 qui était prévu entre septembre 2008 et janvier 2009 a été différée jusqu'en mars 2009. Cette période de latence dans la production de population a été mise à profit pour cribler 181 nouveaux marqueurs microsatellites. Une nouvelle version de la carte génétique, comprenant 277 marqueurs cartographiés sur 21 groupes de liaison avec une taille totale de 1843.7 cM, a été produite. En mars 2009, 593 graines BC_3F_1 putatives représentant les 59 familles BC_2F_1 (5 à 25 graines par plante) ont été semées en serre. Les ADN de 565 individus ont été extraits au CERAAS. A partir de la nouvelle version de la carte génétique, un jeu de 115 marqueurs régulièrement répartis, avec une densité d'environ 1 marqueur tous les 12 cM, a été utilisé pour génotyper les 565 individus. Une première analyse des données génotypiques des 565 individus a permis de sélectionner un sous-ensemble de 284 individus dans lesquels la totalité du génome du parent donneur était représentée et qui avaient un retour au fonds génétique du parent récurrent compris entre 79% et 97%. A partir de ce sous-ensemble, nous avons utilisé le logiciel CSSL-Finder avec les mêmes critères que ceux retenus dans les générations précédentes pour choisir 55 individus BC_3F_1 et 25 BC_2F_2 à avancer en génération suivante. Au total, 80 individus ont été sélectionnés pour produire la génération suivante. Pour des commodités de nomenclature nous considérerons les 80 individus comme étant une même population BC_3F_1 . A l'inverse des générations précédentes, chaque individu a été choisi pour dériver un seul fragment donneur. Ces individus ont été utilisés comme parents femelles pour être croisés avec le parent récurrent et 3020 graines BC_4F_1 ont été récoltées au début du mois de juin 2009.

Choix des individus BC_4F_1 pour développer la population BC_4F_2

Fin juin 2009, 786 graines BC_4F_1 représentant les 80 familles BC_3F_1 (5 à 15 graines par famille) ont été semées en serre au CERAAS. De jeunes feuilles ont été prélevées sur 767 individus BC_4F_1 , séchées à l'étuve pendant 12 heures à 40°C et conservées dans du gel de silice. Les ADN ont été extraits au CIRAD. Le génotypage de ces individus a été réalisé en deux étapes. La première a consisté à confirmer la présence de chaque fragment chromosomique cible issu du donneur à l'aide des marqueurs flanquant. Cette étape a permis

Tableau IV.3 : Résultats de l'analyse des populations F_1 et BC_1F_1 . *estimé avec 200 marqueurs microsatellites, ** estimé sur 12 individus par population

Croisement	N. graines F_1 putatives	N. graines BC_1 putatives	Polymorphisme*	% réussite des croisements**
Fleur11 x AiAd	35	338	57.8	85%
55-437 x AiAd	32	333	57.8	83%
GC8-35 x AiAd	45	386	57.8	83%
73-30 x AiAd	25	86	57.8	100%
GC8-35 x TxAG6	65	<50	72.3	25%

de sélectionner 459 individus qui possédaient les fragments cibles à l'état hétérozygote ou homozygote. La seconde étape a consisté à contrôler le retour au fonds génétique du parent récurrent des 459 individus sélectionnés. Pour chaque individu, les données génotypiques de la génération BC_3F_1 ont été utilisées pour cibler les zones dans lesquelles des fragments donneurs étaient encore présents. Ces zones ont été criblées avec les marqueurs correspondants et 96 individus BC_4F_1 et 4 individus BC_3F_2 possédant en moyenne 4 fragments issus du donneur et un retour au fonds génétique du parent récurrent compris entre 83% et 99% ont été retenus. Pour les mêmes raisons que précédemment nous avons considéré les 100 individus sélectionnés comme étant une même population BC_4F_1 . Contrairement aux générations BC_1F_1 , BC_2F_1 et BC_3F_1 , pour lesquelles le choix d'un nombre minimal d'individus qui offraient une représentation de l'ensemble du génome donneur a été fait pour passer d'une génération à l'autre, en génération BC_4F_1 nous avons gardé autant que possible une certaine redondance (i.e. plusieurs individus sélectionnés pour couvrir une même zone). Les 100 individus BC_4F_1 ont été autofécondés et 6302 graines BC_4F_2 ont été récoltées en Septembre 2009.

Choix des individus BC_4F_2 pour développer la population BC_4F_3

En octobre 2009, 1200 graines BC_4F_2 représentant les 100 familles BC_4F_1 (12 graines par plante) ont été semées en serre au CERAAS. De jeunes feuilles ont été prélevées sur 1180 individus BC_4F_2 , séchées à l'étuve pendant 12 heures à 40°C et conservées dans du gel de silice. Les ADN ont été extraits au CIRAD. Les mêmes étapes que celles décrites pour la génération précédente ont été utilisées pour le génotypage. Le criblage sur le fragment cible a permis de sélectionner un premier sous-ensemble de 877 individus qui présentaient le fragment sous forme homozygote ou hétérozygote. Ce sous-ensemble d'individus a été contrôlé pour le retour sur le fonds génétique du parent récurrent. Pour choisir les individus à avancer en génération BC_4F_3 un tri sur le nombre et la configuration (i.e. homozygote ou hétérozygote) des fragments dans chaque lignée a été réalisé. Les individus homozygotes pour le fragment cible dans un fonds génétique homogène cultivé ont été retenus en premier choix. En second choix nous avons retenu les lignées hétérozygotes pour le fragment cible dans un fonds génétique homogène cultivé ou homozygote pour le fragment cible et ne possédant par ailleurs qu'un seul autre fragment à l'état hétérozygote. Enfin, nous avons choisi les meilleurs individus avec 3, 4 ou 5 fragments qui permettaient de couvrir les zones du génome que les deux premiers cas de figure ne permettaient pas de combler. Au total, 140 individus BC_4F_2 ont été autofécondés et 4326 graines BC_4F_3 ont été récoltées en janvier 2010.

Tableau IV.4 : Caractéristiques des individus des générations successives de backcross

Génération	N. ind. génotypés	N. ind. retenus	% homozygote Fleur 11	% hétérozygote	% homozygote AiAd	taille moyenne fragments (cM)	Min (cM)	Max (cM)	N. moyen de fragments par individu
BC ₁ F ₁	88	22	49.99	49.26	0	40.36	2.33	99.65	17.07
BC ₂ F ₁	192	58	77.16	21.59	0	35.63	2.35	73.92	11
BC ₃ F ₁	565	80	87.31	10.1	1.23	39.23	11.3	115.6	6.43
BC ₄ F ₁	807	100	93.66	6.13	0.24	39.15	6.5	115.6	4.06
BC ₄ F ₂	1180	140	96.17	2.4	1.4	39.5	5.7	115.6	2.12
BC ₄ F ₃	1078	122	97.28	0.14	2.24	39.24	5.7	115.6	1.2

Choix des individus BC₄F₃ à multiplier

En mars 2010, 1100 graines BC₄F₃ représentant les 140 familles BC₄F₂ (3 à 8 graines par plante) ont été semées en serre au CERAAS. De jeunes feuilles ont été prélevées sur 1078 individus BC₄F₃, séchées à l'étuve pendant 12 heures à 40°C et conservées dans du gel de silice. Les ADN ont été extraits au CIRAD. Le criblage sur le fragment cible a permis de sélectionner un sous-ensemble de 718 individus qui présentaient le fragment à l'état homozygote. Le criblage pour le retour au fonds génétique du parent récurrent a permis de sélectionner un sous-ensemble de 122 BC₄F₃. En août 2010, 10 à 20 graines par individu ont été semées en champ à la station de recherche de Bambey pour une multiplication.

III. RESULTATS ET DISCUSSION

III-1 CHOIX DE LA POPULATION D'ETUDE

Les résultats sur le nombre d'individus par population BC₁F₁, le pourcentage de polymorphisme entre parent de chaque population et le taux de réussite des croisements sont consignés dans le Tableau IV.3. La comparaison du pourcentage de polymorphisme entre les parents des différents croisements a montré que le nombre de marqueurs polymorphes était plus important dans le croisement impliquant TxAG6 et GC8-35 (72.3%). Cependant, ce croisement présentait aussi le plus faible nombre d'individus BC₁F₁ (< 50) et le plus faible taux de succès des hybridations (25%). Cette population n'a donc pas été retenue pour la suite des travaux. Le croisement 73-30 x AiAd, bien qu'intéressant en terme de réussite des hybridations avait un nombre limité d'individus BC₁F₁ et était celui pour lequel nous avons trouvé le plus grand nombre d'allèles non-parentaux. De plus 73-30 et AiAd ont tous les deux des graines dormantes ce qui représentait un obstacle majeur pour nos objectifs de réaliser plusieurs générations par an. Cette population a donc aussi été écartée pour la suite des travaux. Le nombre d'individus BC₁F₁ était globalement similaire dans les 3 autres populations impliquant l'amphidiploïde AiAd (i.e. Fleur11 x AiAd, 55-437 x AiAd et GC8-35 x AiAd), avec un taux de réussite des hybridations qui était compris entre 83% et 85%. Ces populations étaient à peu près équivalentes pour les critères relatifs au nombre d'individus BC₁F₁, au pourcentage de polymorphisme et à la réussite des hybridations. Nous avons donc choisi le croisement sur la base des caractères physiologiques et agronomiques des parents récurrents. Le croisement impliquant la variété Fleur11 a été retenu pour les raisons liées au

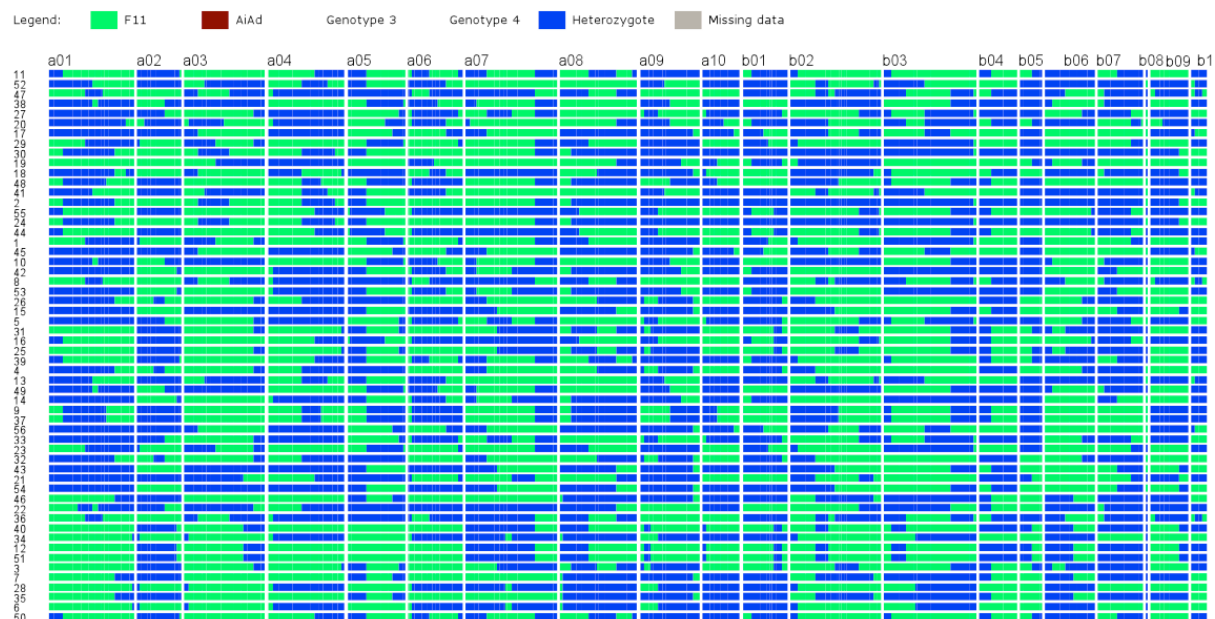


Figure IV.2 : Représentation graphique du génotype des 22 individus BC₁F₁ choisis pour produire la génération BC₂F₁. Vert : homozygote Fleur11, bleu : hétérozygote



Figure IV.3 : Représentation graphique du génotype des 58 individus BC₂F₁ choisis pour produire la génération BC₃F₁. Vert : homozygote Fleur11, bleu : hétérozygote, gris : données manquantes

potentiel de rendement important, à la taille des graines et à l'absence de dormance du parent récurrent.

III-2 DEVELOPPEMENT DE LA COLLECTION DE LIGNEES D'INTROGRESSION

Parmi les 22 individus BC_1F_1 sélectionnés pour développer la génération BC_2F_1 , 12 ont été retenus pour trois fragments cibles, 3 pour 2 fragments et 7 pour un seul fragment. Ces individus avaient en moyenne 49.9% de retour au fonds génétique du parent récurrent, 49.2% d'hétérozygotie, une taille de fragments cibles de 40.3 cM et 17 fragments par individu (Tableau IV.4). Au total, 49 fragments offrant une couverture optimale de l'ensemble des groupes de liaison ont été sélectionnés (Figure IV.2).

En génération BC_2F_1 , les 58 individus sélectionnés avaient en moyenne 77.1% de retour au fonds génétique du parent récurrent, 21.5% d'hétérozygotie, une taille de fragments cibles de 35.6 cM et 11 fragments par individu (Tableau IV.4). De la génération BC_1F_1 à la génération BC_2F_1 nous avons observé une réduction de 56.1% du nombre moyen de locus hétérozygotes et une réduction de 35.5% du nombre moyen de fragments par individu. Les 58 individus BC_2F_1 ont été utilisés pour dériver 65 fragments qui offraient une couverture optimale de l'ensemble des groupes de liaison (Figure IV.3).

Les 80 individus sélectionnés en génération BC_3F_1 présentaient en moyenne 87.3% du génome avec des allèles de Fleur11 à l'état homozygote, 10.1% sous forme hétérozygote et 1.2% avec des allèles d'AiAd à l'état homozygote. Le nombre moyen de fragments par individus était de 6.4 (Tableau IV.4). Entre les générations BC_2F_1 et BC_3F_1 nous avons observé une réduction de 53.2% du nombre moyen de locus hétérozygotes et une réduction de 41.5% du nombre moyen de fragments par individu. L'ensemble des individus BC_3F_1 sélectionnés a permis de couvrir la totalité des groupes de liaison avec des fragments d'une taille moyenne de 39.2 cM (Figure IV.4).

Entre la BC_1F_1 et BC_3F_1 , en dépit du fait que nous avons conduit notre sélection avec une carte non saturée, nous n'avons enregistré aucune perte de fragment chromosomique d'origine sauvage. Cela est principalement lié au choix que nous avons fait de géotyper tous les individus des différentes générations sur l'ensemble du génome. Ce choix, bien que coûteux en terme de points de données à produire (24376 points de données en BC_1F_1 , 16512 en BC_2F_1 et 64975 en BC_3F_1), nous a semblé le plus cohérent dans notre contexte de production des populations.

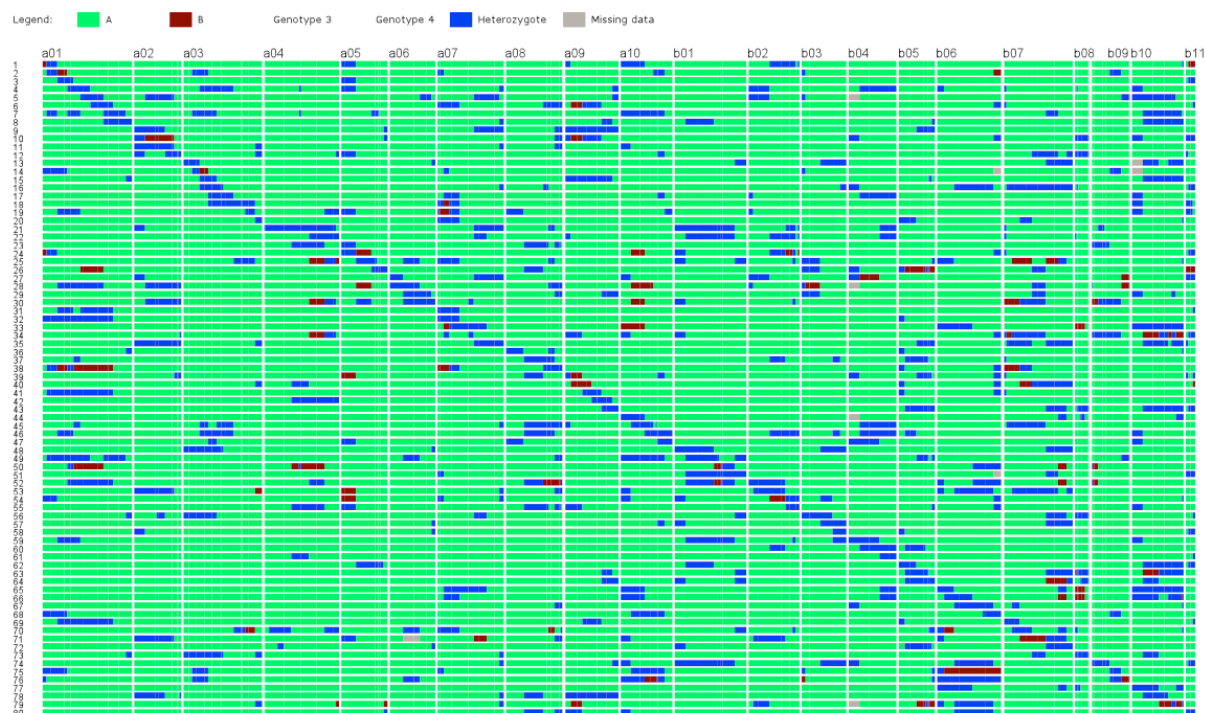


Figure IV.4 : Représentation graphique du génotype des 80 individus BC₃F₁ choisis pour produire la génération BC₄F₁. Vert : homozygote Fleur11, bleu : hétérozygote, bordeaux: homozygote AiAd, gris : données manquantes

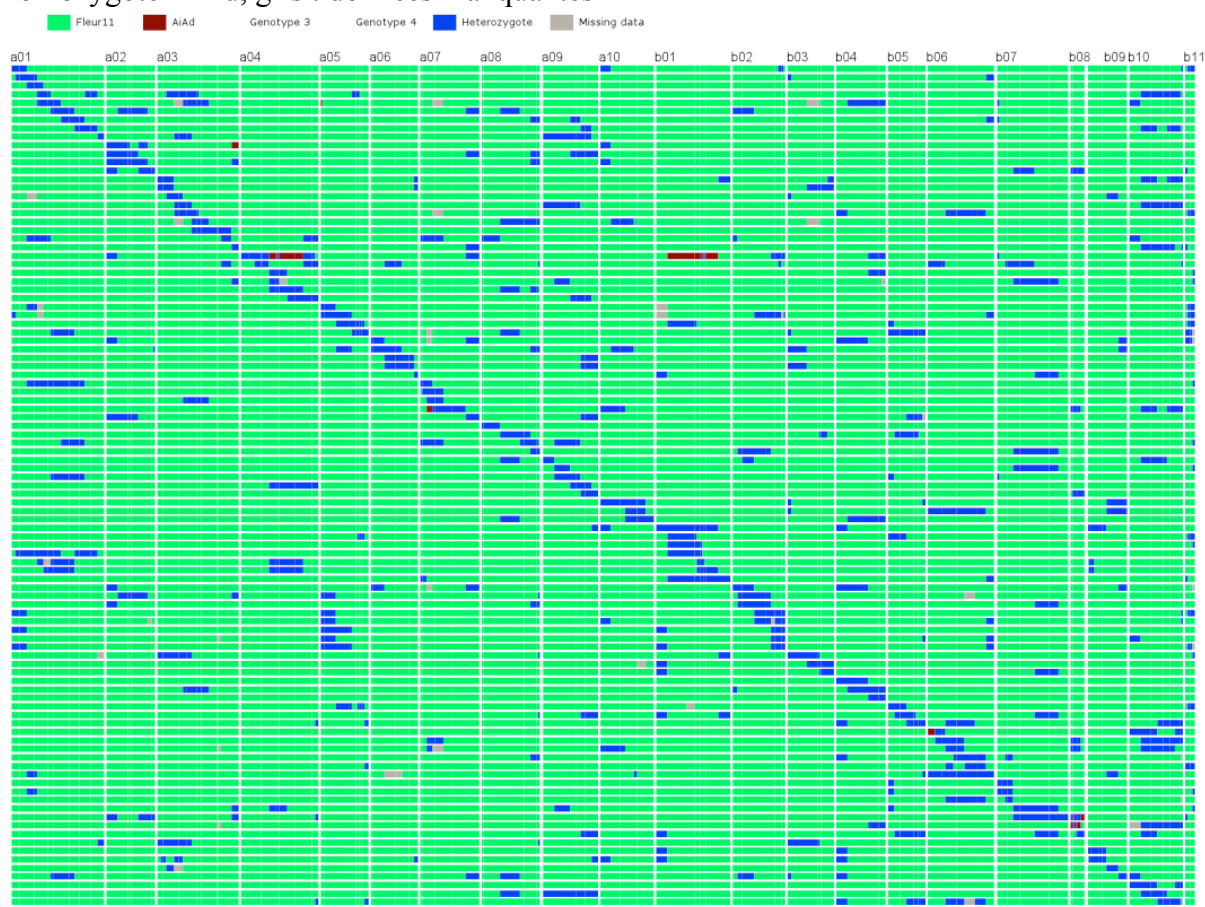


Figure IV.5 : Représentation graphique du génotype des 100 individus BC₄F₁ choisis pour produire la génération BC₄F₂. Vert : homozygote Fleur11, bleu : hétérozygote, bordeaux: homozygote AiAd, gris : données manquantes

Les 100 individus sélectionnés en génération BC₄F₁ présentaient en moyenne 93.6% de retour au fonds génétique du parent récurrent, 6.1% de loci à l'état hétérozygote et 0.24% du génome avec des allèles d'AiAd à l'état homozygote. Le nombre moyen de fragments était de 4.06 par individu (Tableau IV.4). Entre les générations BC₃F₁ et BC₄F₁ nous avons observé une réduction plus faible du nombre moyen de locus hétérozygotes (39.3% contre 56.1% et 53.2% les générations précédentes). Cette faible réduction peut s'expliquer par le fait que la population BC₄F₁ a été dérivée en partie à partir d'individus BC₂F₂. Une réduction de 36.8% du nombre moyen de fragments par individu a aussi été observée. Au total, 20% d'individus ont été choisis pour apporter de la redondance sur certains fragments. L'ensemble des individus BC₄F₁ sélectionnés a permis de couvrir la totalité des groupes de liaison avec des fragments d'une taille moyenne de 39.1 cM (Figure IV.5).

En génération BC₄F₂, 140 individus ayant en moyenne 96.1% de retour au fonds génétique du parent récurrent, 2.4% de locus à l'état hétérozygote et 1.4% du génome avec des allèles d'AiAd à l'état homozygote ont été sélectionnés. Le nombre moyen de fragments par individu étaient de 2.1 (Tableau IV.4). Nous avons enregistré une réduction de 60.8% et de 47.7% du nombre de locus hétérozygotes et du nombre moyen de fragments par individu respectivement. Parmi les 140 individus sélectionnés, 54 (38.5%) ont été choisis pour apporter de la redondance. L'ensemble des individus BC₄F₂ sélectionnés a permis de couvrir la totalité des groupes de liaison avec des fragments d'une taille moyenne de 39.5 cM (Figure IV.6).

Pour la génération BC₄F₃, 122 individus ont été sélectionnés. Ces individus avaient en moyenne 97.2% de retour au fonds génétique du parent récurrent, 0.14% de locus à l'état hétérozygote, 2.24% du génome avec des allèles d'AiAd à l'état homozygote et une taille moyenne de fragments de 39.2 cM (Tableau IV.4). Le nombre moyen de fragments par individu (1.2) a été réduit de 43.3% par rapport à la génération précédente. Parmi les 122 individus sélectionnés, 94 (77%) ne possédaient qu'un fragment du donneur, 23 (18.8%) en possédaient 2, 4 (3.2%) en possédaient 3 et seul un individu avait encore 5 fragments. Tous les fragments du donneur choisis pour couvrir les groupes de liaison étaient fixés et le nombre de locus hétérozygotes a été réduit de 94.1% par rapport à la génération précédente (Figure IV.7). En faisant abstraction de la redondance entre lignées qui portent des fragments choisis pour couvrir la même zone, nous avons identifié un sous-ensemble de 80 lignées qui sont suffisantes pour couvrir la totalité des groupes de liaison.

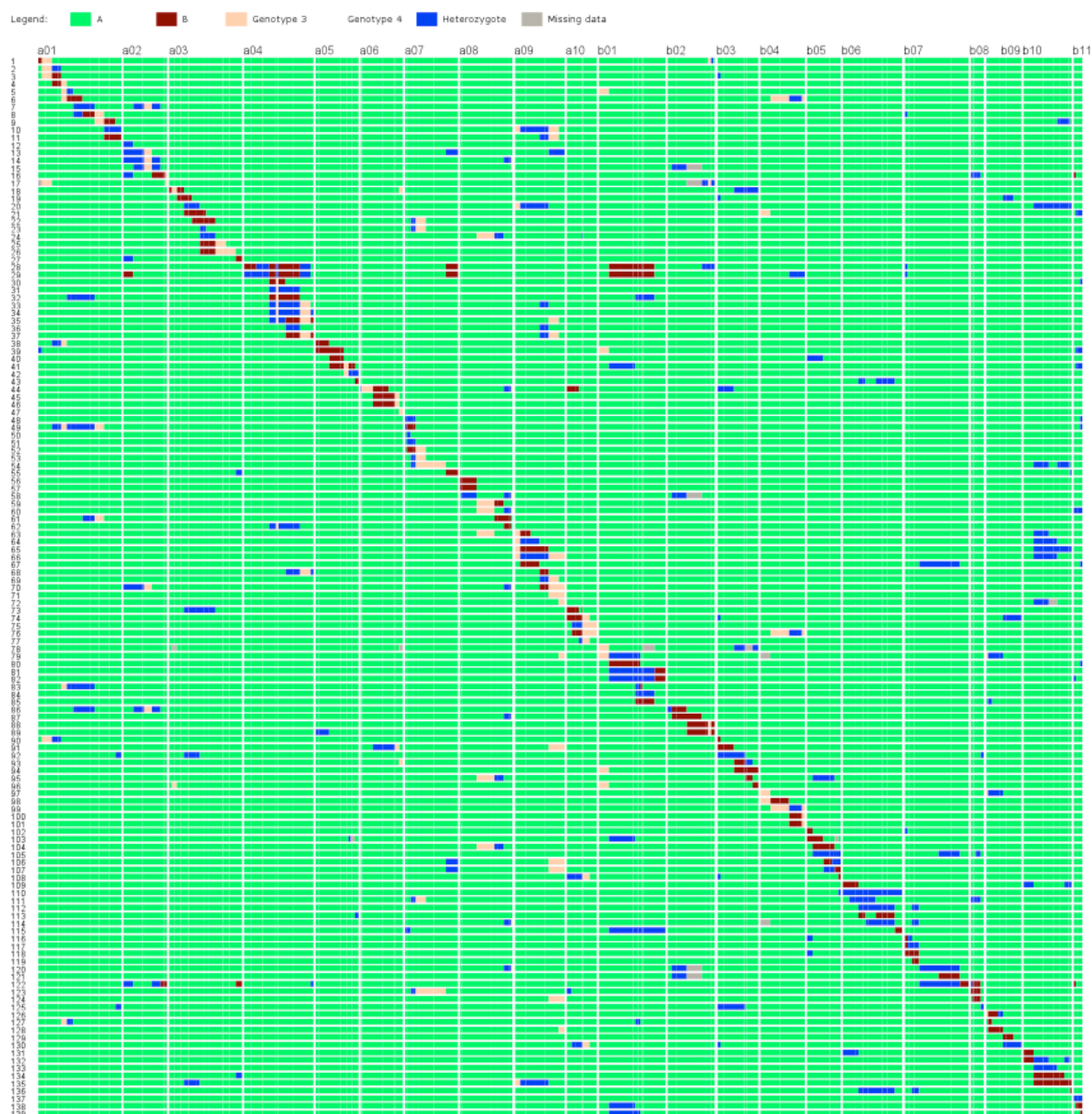


Figure IV.6 : Représentation graphique du génotype des 140 individus BC₄F₂ choisis pour produire la génération BC₄F₃. Vert : homozygote Fleur11, bleu : hétérozygote, bordeaux: homozygote AiAd, saumon : hétérozygote ou homozygote AiAd gris : données manquantes

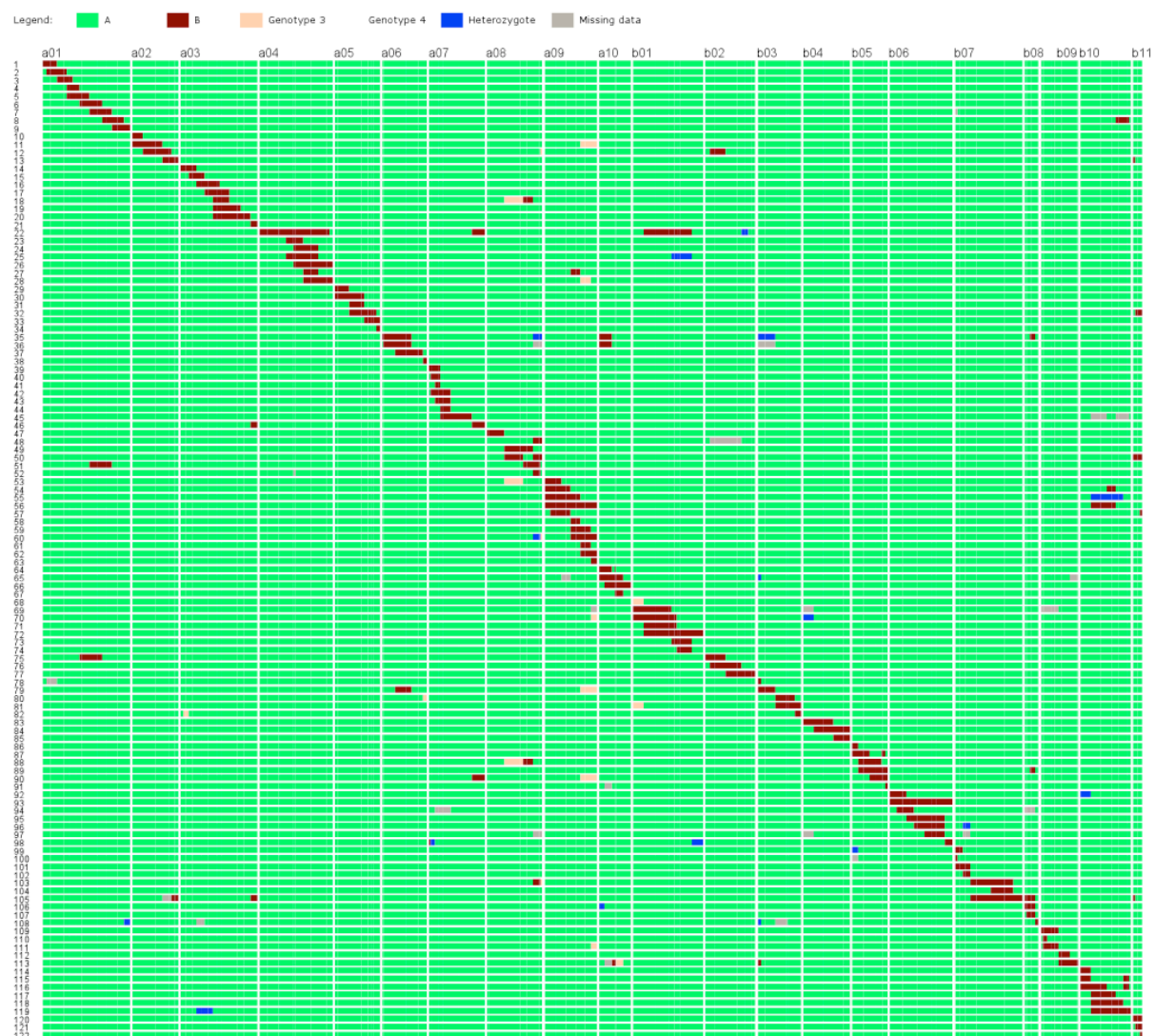


Figure IV.7 : Représentation graphique du génotype des 122 individus BC_4F_3 . Vert : homozygote Fleur11, bleu : hétérozygote, bordeaux: homozygote AiAd, saumon : hétérozygote ou homozygote AiAd gris : données manquantes



Dispositif de conduite des générations au Ceraas

- a. Plantes BC_4F_2 en serre
- b. Plantes BC_4F_2 en extérieur

IV. CONCLUSION

Dans cette étude nous avons développé en trois années, à l'aide d'une stratégie de sélection assistée par marqueurs, une population de 122 lignées d'introgession issue d'un croisement initial entre la variété cultivée Fleur11 et l'amphidiploïde AiAd. Cette population comprend en moyenne 1.2 fragments par lignée, 97.2% de retour au fonds génétique du parent récurrent et une taille moyenne de fragments de 39.2 cM.

La stratégie de génotypage que nous avons mise en place durant les trois premières générations de backcross nous a permis de mener en parallèle la construction des lignées et saturation de la carte génétique sans perdre de fragments issus du donneur. En dépit du nombre limité d'individus que nous avons sélectionné à chaque génération, nous avons obtenu un bon niveau de retour au fonds génétique du parent récurrent en quatre générations de backcross et deux générations d'autofécondation. La redondance que nous avons apportée à partir de la génération BC₄F₁ nous a permis de pallier les éventuelles pertes de fragments et d'optimiser le choix des individus à avancer en génération suivante. L'identification d'un nombre restreint de lignées (80) permettant de couvrir la totalité des groupes de liaison présente un intérêt particulier pour les perspectives d'optimisation des dispositifs expérimentaux pour un phénotypage plus précis de la population.

Cette population de lignées d'introgession est à notre connaissance la première développée chez l'arachide et représente une ressource génétique importante pour la détection de nouveaux QTL, la validation et la cartographie fine de QTL d'intérêt agronomique.

Chapitre V CONCLUSION GENERALE ET PERSPECTIVES

Au cours de cette thèse, une approche originale chez l'arachide a été développée pour utiliser les espèces diploïdes sauvages *A. duranensis* (AA) et *A. ipaensis* (BB) dans l'objectif d'élargir la base génétique de l'espèce cultivée tétraploïde *A. hypogaea* (AABB). Cette approche a d'abord consisté à développer une population BC₁F₁ qui a été utilisée pour construire une carte génétique puis à développer et à caractériser une population AB-QTL pour la cartographie de régions génomiques impliquées dans la domestication de l'espèce cultivée et dans la variation de caractères agronomiques d'intérêts en conditions normale et de déficit hydrique. Une population de lignées d'introgession a ensuite été développée à l'aide d'un schéma de sélection assistée par marqueurs.

Le cadre institutionnel dans lequel cette thèse a été réalisée a mis en exergue des collaborations entre instituts de recherche du Sud (EMBRAPA et ISRA) pour l'échange de matériel génétique et entre instituts de recherche du Sud et du Nord (ISRA et CIRAD) pour le développement de la méthodologie de sélection, la création des populations et pour l'analyse du génotype et du phénotype des individus. Le développement et l'intensification des interactions entre les activités de création des populations au Sénégal et des activités de génotypage en France, matérialisées au travers d'un flux important d'informations et de matériels génétiques, ont été les points forts de la collaboration ISRA-CIRAD autour de ce projet et les clefs de la réussite de la production et de l'analyse des populations d'AB-QTL et de lignées d'introgession.

Ces travaux de thèse ont permis de relever le défi de l'application d'une stratégie de sélection assistée par marqueurs à l'arachide, plante cléistogame à cycle court pour laquelle il est difficile de réaliser des croisements manuels et qui produit peu de descendants par génération de backcross. Pour limiter les risques d'autofécondation inhérents à la biologie de la reproduction de l'arachide, un protocole rigoureux de croisement (émasculature des fleurs et croisement à des horaires précis qui permettent d'optimiser la réussite des hybridations) et un contrôle systématique des descendants à l'aide de marqueurs microsatellites ont été réalisés. Cependant, le handicap que pouvaient représenter les autofécondations au cours du processus de développement des populations de backcross (BC₁F₁ à BC₄F₁) a été transformé en avantage pour la production de la population AB-QTL. En effet, la possibilité de différencier dans une même génération les individus issus d'une autofécondation de ceux issus d'une hybridation a permis de dériver, à partir d'une même plante, les individus de la population

AB-QTL (BC₂F₃ et BC₃F₂) et les individus de la population de lignées d'introgression (BC₃F₁ et BC₄F₁), ce qui a représenté un gain de temps considérable. Le cycle court de la plante (en moyenne 21 jours du semis à la floraison) a contrasté avec le volume important de points de données génotypiques à produire et a nécessité une certaine réactivité dans le processus de production de ces données tant en terme de quantité que de qualité pour choisir les individus à croiser. L'application « Jelly » (Rami, com. pers.) a permis de faciliter l'analyse des images des gels produits au cours du génotypage, de stocker et d'accéder rapidement à ces images et aux données génotypiques afférentes. Cette application a apporté une réelle plus-value dans la vitesse d'analyse des informations génotypiques des différentes générations de backcross.

Les travaux de recherche entrepris au cours de cette thèse ont permis des avancées significatives sur les questions relatives à l'origine et à la domestication de l'espèce cultivée, à la synténie et à l'évolution des génomes A et B de même que leur contribution à l'expression des caractères phénotypiques dans un contexte tétraploïde. Enfin, nous avons identifié des allèles favorables en provenance des espèces sauvages utilisables pour l'amélioration de l'arachide cultivée.

En ce qui concerne l'origine de l'espèce cultivée, les résultats sur l'analyse du polymorphisme des marqueurs microsatellites présentés dans le chapitre 2 ont permis de montrer que les espèces sauvages *A. duranensis* et *A. ipaensis* étaient chacune très proches des génomes A et B de l'espèce cultivée. De plus, le bon niveau de fertilité des individus F₁ du croisement entre Fleur11 et l'amphidiploïde, la ségrégation disomique observée pour l'ensemble des loci cartographiés et le faible pourcentage de marqueurs présentant une distorsion de ségrégation ont été des indicateurs supplémentaires de la forte proximité génétique entre ces espèces sauvages et l'espèce cultivée et ont permis de confirmer *A. duranensis* et *A. ipaensis* comme les plus probables progéniteurs de l'arachide cultivée.

L'analyse des marqueurs communs aux génomes A et B a montré qu'il existe une bonne colinéarité entre ces génomes. Elle a aussi permis de mettre en évidence des réarrangements chromosomiques tels que les inversions de segments entre les groupes de liaison homéologues a01/b01, a03/b03 et a09/b09 ainsi qu'un événement de translocation qui implique les groupes de liaison b07 et b08. Les comparaisons entre la carte produite pendant cette thèse et les cartes génétiques existantes construites à partir d'espèces diploïdes et tétraploïdes ont montré que les réarrangements génomiques observés sont antérieurs à l'événement d'hybridation qui a conduit à la formation de l'espèce cultivée tétraploïde. Ces réarrangements pourraient être impliqués dans la divergence entre les génomes A et B des espèces diploïdes de la section *Arachis*.

En ce qui concerne la domestication de l'espèce cultivée, les résultats sur la cartographie des caractères qui contribuent au syndrome de domestication chez l'arachide, présentés dans le chapitre 3, ont montré que la plupart des QTL impliqués dans l'augmentation de la taille des gousses et des graines sont en cluster sur les groupes de liaison b02, b05 et a07 tandis que ceux impliqués dans la modification de la forme des gousses et du port de la plante sont dispersés sur 11 groupes de liaison. Ces résultats donnent des arguments en faveur d'un processus de domestication initialement ciblé sur la taille des fruits et d'événements ultérieurs de sélection portant sur l'accumulation d'allèles impliqués dans la modification de la morphologie des gousses et du port de la plante. Ces événements ont pu se produire à des endroits et à des périodes différents. Cette évolution spatio-temporelle du processus de fixation des allèles responsables de la modification de la taille des fruits et de la morphologie des gousses et du port de la plante pourrait être à l'origine de la grande diversité morphologique observée dans le compartiment cultivé et qui a prévalu à la subdivision de l'espèce cultivée en deux sous-espèces. Une analyse du processus de domestication en rapport avec la structuration en sous-espèces de l'arachide cultivée permet d'émettre l'hypothèse d'un événement de domestication primaire ayant abouti à des formes primitives proches de celles que l'on retrouve dans la sous-espèce *hypogaea* (port rampant, forte constriction des gousses, graines dormantes de petite taille ...) et des événements de sélection secondaires ayant abouti aux génotypes que l'on retrouve dans la sous-espèce *fastigiata* (port érigé, perte de la constriction des gousses, graines de grosse taille généralement non-dormantes...).

De manière générale, l'analyse AB-QTL a permis d'identifier des régions génomiques d'origine sauvage qui ont un rôle positif sur la variation de caractères d'importance agronomique tels que la date de floraison, la longueur et le diamètre des gousses et des graines, le nombre de gousses et de graines par plante, le pourcentage de gousses mures par plante, le poids de gousses et de fanes ainsi que la biomasse totale. Ces résultats montrent que les espèces sauvages apparentées à l'arachide, bien qu'inférieures à l'espèce cultivée d'un point de vue agronomique, peuvent contribuer positivement à l'amélioration de caractères d'intérêts tant en condition normale qu'en condition de déficit hydrique. Ils ouvrent des voies nouvelles sur l'utilisation des espèces sauvages pour l'amélioration de l'arachide d'autant plus qu'il existe actuellement une véritable dynamique internationale pour la création d'amphidiploïdes par combinaison de différentes espèces diploïdes de génome A et B.

Les résultats de cette thèse ouvrent de nouvelles pistes de recherche sur les aspects relatifs à la saturation de la carte génétique, à l'utilisation des lignées d'introgession pour la détection, la

validation et la cartographie fine de QTL ainsi que pour la création variétale basée sur une stratégie de sélection assistée par marqueurs pour le pyramidage d'allèles favorables d'origine exotique dans un même fonds génétique. Par ailleurs, ces résultats offrent aussi des outils pour aborder des questions de recherche plus fondamentales relatives à l'origine et au processus de domestication de l'espèce cultivée, au clonage des gènes sélectionnés au cours de ce processus et à l'analyse comparative avec les autres légumineuses. Les points de discussion qui suivent portent sur ces différentes pistes de recherche.

I. LA SATURATION DE LA CARTE GENETIQUE.

Comme nous l'avons indiqué dans le chapitre 2, la carte génétique que nous avons produite n'est probablement pas totalement saturée. Ceci est principalement lié au niveau de polymorphisme limité que nous avons observé entre l'amphidiploïde sauvage et l'espèce cultivée, plus particulièrement entre les génomes B des deux parents. La comparaison de notre carte avec les cartes diploïdes (AA) et (BB) publiées par Moretzsohn et al. (2005) (2009) et la carte tétraploïde publiée par Burow et al. (2001) nous permet de penser que notre carte est proche de la saturation pour ce qui concerne les groupes de liaison provenant du génome A mais pas pour ceux provenant du génome B. Le problème de la saturation de la carte génétique se pose avec plus d'acuité dans le contexte de la population de lignées d'introgession pour laquelle l'ensemble du génome de la variété cultivée doit être couvert avec un fragment du génome du parent sauvage. Un effort de cartographie de nouveaux marqueurs est donc nécessaire pour identifier les zones du génome de la variété cultivée qui ne seraient pas couvertes par un fragment du donneur sauvage, et dans une moindre mesure, les fragments du donneur qui seraient encore présents dans la population. Un second objectif de la cartographie de nouveaux marqueurs est l'augmentation de la densité de la carte pour les applications de cartographie fine de QTL.

Dans l'objectif de saturation de la carte génétique, nous avons, en partenariat avec l'équipe australienne DArT P/L, génotypé les 88 individus la population BC₁F₁ à l'aide de marqueurs DArT. Les premiers résultats du génotypage de ces individus étaient très mitigés avec environ 50 marqueurs polymorphes sur les 1800 clones de la puce DArT. De plus, les 50 marqueurs polymorphes étaient fortement redondants limitant ainsi le nombre de marqueurs utilisables pour la cartographie génétique. Le faible polymorphisme des marqueurs DArT s'expliquait par une faible représentation de génotypes sauvages lors du développement de la puce. Le développement d'une puce DArT qui est mieux adaptée à notre problématique de

cartographie génétique a été réalisé au CIRAD. Sur cette nouvelle puce sont représentés 3000 séquences d'ADN de chaque espèce diploïde sauvage *A. duranensis* et *A. ipaensis*, soit au total une puce contenant 6000 clones. La cartographie de marqueurs DArT à partir de cette nouvelle puce est en cours au CIRAD. Ces marqueurs permettront d'avoir une meilleure couverture du génome.

A l'heure actuelle, environ 3000 marqueurs microsatellites, développés à partir de banques d'EST, de banques BAC et de banques enrichies en microsatellites, sont disponibles chez l'arachide. Il n'existe pas encore d'informations sur le niveau de redondance qui existe entre les différentes séquences à partir desquelles ces marqueurs ont été produits. Cependant une prévision optimiste basée sur la faible redondance des marqueurs microsatellites et sur le niveau de polymorphisme qui existe entre les génomes B sauvage/cultivé (~30%) permet d'estimer à quelques centaines le nombre de marqueurs supplémentaires potentiellement utilisables pour la saturation de la carte.

Enfin, la saturation de la carte génétique à l'aide de marqueurs SNP est envisagée dans le cadre de la collaboration que nous avons initiée avec une équipe de l'UGA (The University of Georgia). Cette équipe a développé 1536 marqueurs SNP à partir de séquences cDNA de deux accessions de l'espèce diploïde *A. duranensis*. Ces SNP ont été cartographiés sur une population F₂ issue du croisement entre ces deux accessions (Knapp et al. en préparation). Cependant, l'utilisation de cette ressource SNP dans un contexte allotétraploïde est soumise à quelques réserves techniques relatives aux difficultés d'interprétations du polymorphisme SNP entre génomes homologues et homéologues. Au-delà de ces défis techniques, il se pose aussi la question de l'efficacité d'une ressource SNP développée à partir de séquences du génome A pour révéler du polymorphisme sur le génome B. Néanmoins, ces marqueurs SNP représentent une perspective intéressante pour les objectifs de saturation de la carte génétique.

II. CARACTERISATION DE LA POPULATION DE LIGNEES D'INTROGRESSION.

Une caractérisation moléculaire de la population de lignées d'introgession, à l'aide des nouveaux marqueurs qui seront placés sur la carte, devra être effectuée afin d'identifier les zones du génome qui ne seraient pas couvertes de même que les fragments surnuméraires du parent donneur. La recherche des fragments du parent donneur qui seraient absents dans la population pourra se faire par génotypage des individus des générations précédentes (BC₄F₂, BC₄F₁, BC₃F₁ et BC₂F₁) et rétrocroisement des individus détenteurs du(es) fragment(s) avec

le parent récurrent. De même, il est à envisager un rétrocroisement des lignées possédant des fragments surnuméraires du parent donneur pour produire des lignées à introgression unique. A court terme, l'une des principales suites à donner aux travaux de cette thèse porte sur la caractérisation phénotypique de la population de lignées d'introgression pour la validation des QTL détectés lors de l'analyse AB-QTL (chapitre 3) et pour l'identification de nouveaux QTL. Une analyse phénotypique de cette population pour plusieurs caractères et dans plusieurs environnements est prévue dans le cadre d'un nouveau projet intitulé « Tropical Legume I phase 2 » (TL1), d'une durée de quatre ans, financé par la fondation Bill et Melinda Gates. Dans ce projet, le jeu de 80 lignées d'introgression sera caractérisé dans une approche de « phénotypage à la carte » qui tire profit du caractère pérenne de la population et des compétences des différents partenaires. En ce qui concerne la validation des QTL et la détection de nouveaux QTL impliqués dans le rendement en conditions normale et de déficit hydrique, une évaluation de la population sera faite dans trois environnements par les partenaires du Sénégal, du Mali et du Niger. Une évaluation plus fine des caractères physiologiques explicatifs des mécanismes mis en jeu par la plante pour la tolérance à la sécheresse et la cartographie de QTL afférents sera effectuée par les partenaires de l'ICRISAT (Inde). Toujours dans le cadre de ce projet, une évaluation des lignées pour la résistance aux maladies foliaires sera effectuée par les partenaires de l'EMBRAPA (Brésil). Une caractérisation de la population d'ILs pour l'étude des interactions entre la tolérance à la sécheresse et la contamination à l'aflatoxine au champ est envisagée dans le cadre d'un projet BREAD en cours de montage avec les partenaires de l'UGA, l'ICRISAT et l'EMBRAPA. Les phénotypages multi-environnements pour la tolérance à la sécheresse permettront de valider les QTL détectés précédemment, d'estimer les interactions génotype x environnement et d'identifier les QTL stables dans différents environnements ainsi que les QTL spécifiques d'un environnement donné. L'analyse multi-caractère (physiologie de la sécheresse, résistance aux maladies foliaires et à la contamination à l'aflatoxine) permettra d'identifier de nouvelles régions génomiques d'origine sauvage associées à la variation positive de ces caractères.

Enfin, afin de d'exploiter la large gamme de variabilité existante dans la population d'ILs, il serait intéressant d'étendre le phénotypage à d'autres caractères tels que la qualité des huiles (sélection d'allèles des parents sauvages qui améliorent le ratio O/L), la qualité fourragère des fanes, la résistance à la rouille, à la maladie de la rosette et aux nématodes...

III. CREATION VARIETALE PAR PYRAMIDAGE DE QTL ASSISTEE PAR MARQUEURS.

Au travers de l'analyse AB-QTL, plusieurs régions génomiques d'origine sauvage se sont révélées intéressantes pour l'amélioration de caractères impliqués dans les composantes du rendement. La particularité de ces régions porte sur une colocalisation de plusieurs QTL impliqués dans différentes composantes du rendement en conditions normale et/ou de déficit hydrique et, dans la limite de détection de QTL de notre expérimentation, sur une absence de caractères indésirables. Il s'agit par exemple :

- d'une région d'environ 12 cM sur le groupe de liaison a01 entre les marqueurs TC2E05_A et IPAHM287_A sur laquelle sont cartographiés des QTL de poids et de nombre de gousses par plante,
- d'une région d'environ 30 cM sur le groupe de liaison a05 entre les marqueurs gi-0620_A et Ah-614_A dans laquelle colocalisent des QTL de nombre et de poids de graines par plante ainsi que de 5 QTL relatifs aux indices de tolérance à la sécheresse,
- d'une région d'environ 20 cM sur le groupe de liaison a10 entre les marqueurs TC1G04_A et AC2B03_A qui porte des QTL de taille de gousses,
- et d'une région d'environ 43 cM sur le groupe de liaison b06 entre les marqueurs PM137_B et TC11A04_B sur laquelle on retrouve des QTL de taille de gousses et de graines.

Dans la population d'ILs il existe quatre lignées qui ont une introgression unique de l'un de ces fragments. Après validation de ces QTL, ces lignées peuvent constituer le point de départ d'un programme de création variétale par pyramidage de ces quatre régions chromosomiques dans un même idéotype. Ces lignées peuvent aussi être utilisées pour développer des NILs, par croisement avec le parent récurrent, qui permettent de réduire la taille du fragment à introgresser. Cependant le développement de NILs peut représenter un investissement important.

D'autres régions génomiques d'origine sauvage portent des QTL intéressants d'un point de vue agronomique qui colocalisent cependant avec des QTL impliqués dans l'expression de caractères indésirables tels qu'une forte constriction des gousses ou un bec proéminent. Il s'agit par exemple :

- d'une région d'environ 20 cM sur le groupe de liaison a08 entre les marqueurs TC22C01_A et IPAHM229_A qui porte un QTL de longueur de gousses et un QTL de constriction de gousses,

- d'une grande région d'environ 60 cM sur le groupe de liaison a09 entre les marqueurs TC9B07_A et PM170_A qui porte des QTL de longueur de gousses et de graines qui colocalisent avec un QTL de constriction des gousses et un QTL de forme de bec
- et d'une région de 15 cM sur le groupe de liaison b11 entre les marqueurs TC3E02_B et TC23E04_B dans laquelle colocalisent des QTL de pourcentage de maturité, de constriction et de bec des gousses.

Dans la population d'ILs, il existe plusieurs lignées consécutives à fragments chevauchants qui couvrent chacune des régions et qui pourraient être utilisées pour affiner la localisation des QTL à effets favorables. Ces lignées pourront ensuite être utilisées pour développer des NILs dans l'objectif de dissocier les caractères positifs des caractères indésirables et d'utiliser les meilleures NILs comme parent de départ pour le pyramidage des QTL.

IV. ORIGINE ET DOMESTICATION DE L'ESPECE CULTIVEE

En l'état actuel des connaissances, l'origine monophylétique de l'arachide cultivée est globalement admise par la communauté de chercheurs qui travaillent sur l'arachide. Cependant, il se pose encore la question des mécanismes qui ont produit, à partir d'un seul événement d'hybridation entre deux espèces sauvages suivi d'un doublement du stock chromosomique, la large diversité morphologique observée dans le compartiment cultivé. Un des résultats les plus frappants lors de l'analyse de la population AB-QTL est la grande variabilité des caractères morphologiques qui existait entre les familles BC₂F₂ et BC₃F₁. Pour le port de la plante par exemple, nous avons observé une gamme de variabilité allant d'une stature prostrée à totalement érigée avec un grand nombre de formes intermédiaires portant plus ou moins de branches secondaires et tertiaires. Les mêmes observations ont été faites pour les caractères relatifs à la constriction et à la forme du bec des gousses, à la forme et à la couleur des feuilles et à la présence ou l'absence de fleurs sur l'axe principal. En se référant aux descripteurs agro-morphologiques pour l'arachide publiés par l'IBPGR et l'ICRISAT (1992), pour la plupart des caractères mentionnés ci-dessus, la gamme de variabilité présente dans la population AB-QTL recouvrait celle observée dans le compartiment cultivé. Ces observations suggèrent qu'il serait possible de reproduire des types morphologiques similaires à ceux que l'on retrouve dans les deux sous-espèces de l'arachide cultivée à partir du croisement entre une variété cultivée à port érigé et l'amphidiploïde sauvage combinant les génomes diploïdes A et B des plus probables ancêtres de l'espèce domestiquée. Des régions des génomes sauvages impliquées dans la modification de la plupart des caractères

morphologiques ont déjà été identifiées et des ILs portant des fragments recouvrant ces zones ont aussi été produites. Une reconstitution des morphotypes cultivés par combinaison de fragments d'origine sauvage portés par différentes ILs serait un argument supplémentaire en faveur d'une origine monophylétique de l'arachide cultivée et permettrait de confirmer notre hypothèse d'un événement de domestication primaire ayant produit des formes primitives proches de la sous-espèce *hypogaea* et des événements de diversification par un processus de fixation de nouveaux allèles ayant produit des formes plus évoluées proches de la sous-espèce *fastigiata*. Les formes intermédiaires pourraient résulter d'un flux de gènes entre les deux sous-espèces.

V. CARTOGRAPHIE FINE ET CLONAGE DES GENES SOUS-JACENTS AUX QTL

Les lignées d'introgression sont un excellent matériel de départ pour la cartographie fine de QTL et le clonage de gènes. Elles ont été utilisées avec succès chez la tomate pour la cartographie et le clonage du gène *fw2.2* responsable de la taille des fruits (Alpert and Tanksley, 1996; Frary et al., 2000) et chez le riz pour cloner les gènes *Hd1* et *Hd6* impliqués dans la précocité (Yano et al., 2001; Takahashi et al., 2001). Chez ces deux espèces, l'identification des gènes sous-jacents aux QTL a été réalisée à partir d'approches combinant des cartographies à haute résolution génétique à l'aide de NILS et physique à l'aide de YAC (Yeast Artificial Chromosome).

Chez l'arachide, nous disposons actuellement des ressources génétiques et génomiques de base nécessaires pour entreprendre le clonage de gènes sous-jacents au QTL. En effet, deux banques BAC des espèces diploïdes sauvages *A. duranensis* (génome AA) et *A. ipaensis* (génome BB) ont été récemment construites au CIRAD (Guimarães et al., 2008). La banque BAC du génome A représente 7.4 génome équivalent avec environ 84.000 clones contenant des inserts d'une taille moyenne de 110 Kb. Celle du génome B est de 5.3 génome équivalent avec environ 75.500 clones contenant des inserts d'une taille moyenne de 100 Kb. La banque BAC du génome A a été récemment étendue et le contigage des BAC ainsi que leur alignement sur une carte génétique diploïde ont été effectués par une équipe de l'UGA (Paterson et al., en préparation). Par ailleurs, des marqueurs COS (Conserved Orthologous Sequence) ont été développés et cartographiés chez plusieurs légumineuses dont l'arachide et ont permis de révéler un bon niveau de synténie entre les légumineuses (Hougaard et al., 2008; Bertoli et al., 2009). Ceci donne l'accès aux séquences orthologues des espèces modèles *Medicago truncatula* et *Lotus japonicus*.

Du point de vue des ressources génétiques, les lignées d'introgression produites portent en majeure partie une introgression unique. Certaines ILs, portant des fragments qui couvrent des régions sur lesquelles des QTL ont été identifiés, peuvent être dès à présent utilisées pour développer des NILs. C'est le cas par exemple de six ILs consécutives, portant des fragments chevauchants, choisies pour couvrir une région de 21 cM en position proximale sur le groupe de liaison a07 sur laquelle a été identifiée un cluster de QTL impliqués dans la taille des gousses et des graines. Un phénotypage de ces ILs permettrait d'affiner la localisation du(es) QTL(s) et d'identifier un nombre plus restreint de lignées (1 ou 2) à partir desquelles pourront être développées de larges populations F₂ par croisement avec le parent récurrent. De nouveaux marqueurs moléculaires pourront être développés et cartographiés sur la population F₂ pour augmenter la densité de marquage de cette zone et un génotypage de la population F₂ permettrait d'identifier les individus ayant recombinés autour du QTL. Un phénotypage de ces individus permettrait d'augmenter la résolution autour du gène. Les informations sur la cartographie physique des clones BAC pourraient être utilisées pour identifier des contigs de clones BAC qui couvrent les régions qui ont recombiné autour du gène chez les individus F₂. Les séquences des extrémités des BACs ou des BAC entiers pourraient permettre de développer de nouveaux marqueurs pour restreindre à quelques dizaines de kb la région du gène et de l'identifier.

Les ressources génétiques (collection d'ILs) et génomiques (banque BAC) qui existe actuellement chez l'arachide montrent qu'il est possible de combiner des approches de cartographie génétique et physique pour le clonage de gènes sous-jacents au QTL

VI. VERS LA CONSTRUCTION DE LARGES COLLECTIONS DE LIGNEES D'INTROGRESSION

A l'instar des travaux d'élargissement de la base génétique chez la tomate et chez le riz (Grandillo et al., 2007), il pourrait être mis en place chez l'arachide cultivée un programme systématique d'introgression d'allèles exotiques issus du pool de gènes secondaires. Ce programme, sous réserve de la production d'amphidiploïdes fertiles et compatibles avec l'espèce cultivée, ouvrirait l'accès à la diversité existante dans le pool génique secondaire de l'arachide cultivée, ce qui représente environ une vingtaine d'espèces sauvages. A l'heure actuelle, quelques dizaines d'amphidiploïdes combinant des génomes d'espèces différentes ont déjà été produits par les partenaires de l'ICRISAT et de l'EMBRAPA. Certains d'entre

eux ont montré une bonne fertilité et sont compatibles avec l'espèce cultivée. D'autres sont très peu fertiles et seront difficilement utilisables en croisement avec l'espèce cultivée.

Bien que des efforts supplémentaires de caractérisation des espèces diploïdes et d'identification des couples d'espèces à même de produire les amphidiploïdes les plus fertiles soient nécessaires pour avoir un plus large accès au compartiment sauvage, la voie est ouverte pour une utilisation plus efficace du pool génique secondaire de l'arachide cultivée.

REFERENCES

- ALBERT, C.M., J.M. GAZIANO, W.C. WILLETT, AND J.A. MANSON. 2002. Nut consumption and decreased risk of sudden cardiac death in the Physicians' Health Study. *Archives of Internal medicine* 162: 1382.
- ALPERT, K., AND S. TANKSLEY. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing fw2.2: A major fruit weight quantitative trait locus in tomato. *Proceedings of the National Academy of Sciences of the United States of America* 93: 15503-15507.
- ANDERSON, W.F., C.C. HOLBROOK, AND T.B. BRENNEMAN. 1993. Resistance to *Cercosporidium personatum* Within Peanut Germplasm 1. *Peanut Science* 20: 53-57.
- ASHIKARI, M., AND M. MATSUOKA. 2006. Identification, isolation and pyramiding of quantitative trait loci for rice breeding. *Trends in Plant Science* 11: 344-350.
- BERNACCHI, D., T. BECK-BUNN, D. EMMATTY, Y. ESHED, S. INAI, J. LOPEZ, V. PETIARD, H. SAYAMA, J. UHLIG, D. ZAMIR, AND S. TANKSLEY. 1998. Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *TAG Theoretical and Applied Genetics* 97: 170-180.
- BERNACCHI, D., T. BECK-BUNN, Y. ESHED, J. LOPEZ, V. PETIARD, J. UHLIG, D. ZAMIR, AND S. TANKSLEY. 1998. Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *TAG Theoretical and Applied Genetics* 97: 381-397.
- BERTIOLI, D., M. MORETZSOHN, L. MADSEN, N. SANDAL, S. LEAL-BERTIOLI, P. GUIMARAES, B. HOUGAARD, J. FREDSLUND, L. SCHAUER, A. NIELSEN, S. SATO, S. TABATA, S. CANNON, AND J. STOUGAARD. 2009. An analysis of synteny of *Arachis* with *Lotus* and *Medicago* sheds new light on the structure, stability and evolution of legume genomes. *BMC Genomics* 10: 45.
- BIANCHI-HALL, C.M., R.D. KEYS, H.T. STALKER, AND J.P. MURPHY. 1993. Diversity of seed storage protein patterns in wild peanut (*Arachis*, *Fabaceae*) species. *Plant Systematics and Evolution* 186: 1-15.
- BLAIR, M.W., G. IRIARTE, AND S. BEEBE. 2006. QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean x wild common bean (*Phaseolus vulgaris* L.) cross. *TAG. Theoretical and Applied Genetics* 112: 1149-1163.
- BOONSAENG, T., AND S.M. FLETCHER. 2008. European Union Import Demand for In-Shell Peanuts. *Journal of Agricultural and Applied Economics* 40: 941-951.
- BRIEND, A. 2001. Highly Nutrient-Dense Spreads: A New Approach to Delivering Multiple Micronutrients to High-Risk Groups. *British Journal of Nutrition* 85: S175-S179.

- BUROW, M.D., C.E. SIMPSON, J.L. STARR, AND A.H. PATERSON. 2001. Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.). broadening the gene pool of a monophyletic polyploid species. *Genetics* 159: 823-837.
- BUROW, M.D., C.E. SIMPSON, M.W. FARIES, J.L. STARR, AND A.H. PATERSON. 2009. Molecular biogeographic study of recently described B- and A-genome *Arachis* species, also providing new insights into the origins of cultivated peanut. *Genome* 52: 107-119.
- CANADY, M.A., V. MEGLIC, AND R.T. CHETELAT. 2005. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome* 48: 685-697.
- CHETELAT, R.T., AND V. MEGLIC. 2000. Molecular mapping of chromosome segments introgressed from *Solanum lycopersicoides* into cultivated tomato (*Lycopersicon esculentum*). *TAG Theoretical and Applied Genetics* 100: 232-241.
- CLAVEL, D., P. BARADAT, J.L. KHALFAOUI, N.K. DRAME, N.D. DIOP, O. DIOUF, AND Y. ZUILY-FODIL. 2005. Adaptation à la sécheresse et création variétale: le cas de l'arachide en zone sahélienne. *Oléagineux, Corps Gras, Lipides* 12: 248-60.
- CLAVEL, D., P. BARADAT, J.L. KHALFAOUI, N.K. DRAME, N. DIOP, O. DIOUF, AND Y. ZUILY-FODIL. 2007. Adaptation à la sécheresse et création variétale: le cas de l'arachide en zone sahélienne: Deuxième partie: une approche pluridisciplinaire pour la création variétale. *Oléagineux, corps gras, lipides* 14: 293-308.
- COMPENDIUM OF PEANUT DISEASE. 1997. The American Phytopathological Society. N. Kokalis-Burelle, D. M. Porter, R. Rodriguez-kabana, D. H. Smith, P. Subrahmanyam, St. Paul, Minnesota 55121-2097 USA.
- CUC, L., E. MACE, J. CROUCH, V. QUANG, T. LONG, AND R. VARSHNEY. 2008. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). *BMC Plant Biology* 8: 55.
- DWIVEDI, S.L., S. PANDE, J.N. RAO, AND S.N. NIGAM. 2002. Components of resistance to late leaf spot and rust among interspecific derivatives and their significance in a foliar disease resistance breeding in groundnut (*Arachis hypogaea* L.). *Euphytica* 125: 81-88.
- ESHED, Y., AND D. ZAMIR. 1996. Less-Than-Additive Epistatic Interactions of Quantitative Trait Loci in Tomato. *Genetics* 143: 1807-1817.
- FARQUHAR, G., M. O'LEARY, AND J. BERRY. 1982. On the Relationship Between Carbon Isotope Discrimination and the Intercellular Carbon Dioxide Concentration in Leaves. *Functional Plant Biology* 9: 121-137.
- FAVERO, A.P., C.E. SIMPSON, J.F.M. VALLS, AND N.A. VELLO. 2006. Study of the Evolution of Cultivated Peanut through Crossability Studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. *Crop Science* 46: 1546-1552.

- FERGUSON, M., M. BUROW, S. SCHULZE, P. BRAMEL, A. PATERSON, S. KRESOVICH, AND S. MITCHELL. 2004. Microsatellite identification and characterization in peanut (*A. hypogaea* L.). *TAG Theoretical and Applied Genetics* 108: 1064-1070.
- FERGUSON, M.E., A. JARVIS, H.T. STALKER, D.E. WILLIAMS, L. GUARINO, J.F. VALLS, R.N. PITTMAN, C.E. SIMPSON, AND P.J. BRAMEL. 2005. Biogeography of wild *Arachis* (*Leguminosae*): distribution and environmental characterisation. *Biodiversity and Conservation* 14: 1777-1798.
- FERNÁNDEZ, A., AND A. KRAPOVICKAS. 1994. Cromosomas y evolución en *Arachis* (*Leguminosae*). *Bonplandia* 8: 187-220.
- FERNANDEZ, C. 1993. Effective selection criteria for assessing plant stress tolerance. In *Adaptation of Food Crops to Temperature and Water Stress*, 257-270. C. G. Kuo, Taiwan.
- FISCHER, R.A. 1981. Optimizing the use of water and nitrogen through breeding of crops. *Plant and Soil* 58: 249-278.
- FISCHER, R., AND R. MAURER. 1978. Drought resistance in spring wheat cultivars. I. Grain yield responses. *Australian Journal of Agricultural Research* 29: 897-912.
- FLETCHER, S., AND D. NADOLNYAK. 2006. Strategic Behavior and Trade in Agricultural Commodities B-Competition in World Peanut Markets. In 2006 Annual Meeting, August 12-18, 2006, Queensland, Australia,
- FRARY, A., S. DOGANLAR, A. FRAMPTON, T. FULTON, J. UHLIG, H. YATES, AND S. TANKSLEY. 2003. Fine mapping of quantitative trait loci for improved fruit characteristics from *Lycopersicon chmielewskii* chromosome 1. *Genome* 46: 235-243.
- FRARY, A., L. FRITZ, AND S. TANKSLEY. 2004. A comparative study of the genetic bases of natural variation in tomato leaf, sepal, and petal morphology. *Theoretical and Applied Genetics* 109: 523-533.
- FRARY, A., T.M. FULTON, D. ZAMIR, AND S.D. TANKSLEY. 2004. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *L. pennellii* cross and identification of possible orthologs in the Solanaceae. *TAG Theoretical and Applied Genetics* 108: 485-496.
- FRARY, A., T.C. NESBITT, A. FRARY, S. GRANDILLO, E.V.D. KNAAP, B. CONG, J. LIU, J. MELLER, R. ELBER, K.B. ALPERT, AND S.D. TANKSLEY. 2000. *fw2.2*: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. *Science* 289: 85-88.
- FRASER, G.E. 2000. Nut consumption, lipids, and risk of a coronary event. *Asia Pacific Journal of Clinical Nutrition* 9: S28-S32.
- FRIDMAN, E., F. CARRARI, Y. LIU, A.R. FERNIE, AND D. ZAMIR. 2004. Zooming In on a Quantitative Trait for Tomato Yield Using Interspecific Introgressions. *Science* 305: 1786-1789.

- FULTON, T.M., T. BECK-BUNN, D. EMMATTY, Y. ESHED, J. LOPEZ, V. PETIARD, J. UHLIG, D. ZAMIR, AND S.D. TANKSLEY. 1997. QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *TAG Theoretical and Applied Genetics* 95: 881-894.
- FULTON, T.M., S. GRANDILLO, T. BECK-BUNN, E. FRIDMAN, A. FRAMPTON, J. LOPEZ, V. PETIARD, J. UHLIG, D. ZAMIR, AND S.D. TANKSLEY. 2000. Advanced backcross QTL analysis of a *Lycopersicon esculentum* x *Lycopersicon parviflorum* cross. *TAG Theoretical and Applied Genetics* 100: 1025–1042.
- GARCIA, G.M., S.P. TALLURY, H.T. STALKER, AND G. KOCHERT. 2006. Molecular analysis of *Arachis* interspecific hybrids. *TAG Theoretical and Applied Genetics* 112: 1342-1348.
- GIMENES, M.A., C.R. LOPES, M.L. GALGARO, J.F.M. VALLS, AND G. KOCHERT. 2002. RFLP analysis of genetic variation in species of section *Arachis*, genus *Arachis* (*Leguminosae*). *Euphytica* 123: 421 - 429.
- GIMENES, M., A. HOSHINO, A. BARBOSA, D. PALMIERI, AND C. LOPES. 2007. Characterization and transferability of microsatellite markers of the cultivated peanut (*Arachis hypogaea*). *BMC Plant Biology* 7: .
- GONG, Y., A. HOUNSA, S. EGAL, P.C. TURNER, A.E. SUTCLIFFE, A.J. HALL, K. CARDWELL, AND C.P. WILD. 2004. Postweaning exposure to aflatoxin results in impaired child growth: a longitudinal study in Benin, West Africa. *Environmental health perspectives* 112: 1334.
- GRANDILLO, S., AND S. TANKSLEY. 2005. Advanced backcross QTL analysis: results and perspectives. In: Tuberosa R, Phillips RL, Gale M (eds) *The wake of the double helix: from the green revolution to the gene revolution*, PP 115-132. Bologna.
- GRANDILLO, S., S.D. TANKSLEY, AND D. ZAMIR. 2007. Exploitation of natural biodiversity through genomics. *Genomics-Assisted Crop Improvement* 121–150.
- GREEN, C.C., AND J.C. WYNNE. 1986. Diallel and generation means analyses for the components of resistance to *Cercospora arachidicola* in peanut. *TAG Theoretical and Applied Genetics* 73: 228–235.
- GREGORY, M.P., AND W.C. GREGORY. 1979. Exotic germ plasm of *Arachis* L. interspecific hybrids. *Journal of Heredity* 70: 185-193.
- GREGORY, W.C., AND M.P. GREGORY. 1976a. *Arachis hypogaea* (*Leguminosae-Papilionadae*). *Evolution of crop plants* 134: 151.
- GREGORY, W.C., AND M.P. GREGORY. 1976b. *Arachis hypogaea* (*Leguminosae-Papilionatae*). *Evolution of crop plants* 134: 151.
- GRIEL, A.E., B. EISSENSTAT, V. JUTURU, G. HSIEH, AND P.M. KRIS-ETHERTON. 2004. Improved diet quality with peanut consumption. *Journal of the American College of Nutrition* 23: 660-668.

- GUIMARÃES, P.M., O. GARSMEUR, K. PROITE, S.C.M. LEAL-BERTIOLI, G. SEIJO, C. CHAINE, D.J. BERTIOLI, AND A. D'HONT. 2008. BAC libraries construction from the ancestral diploid genomes of the allotetraploid cultivated peanut. *BMC plant biology* 8: 14.
- GUR, A., AND D. ZAMIR. 2004. Unused Natural Variation Can Lift Yield Barriers in Plant Breeding. *PLoS Biol* 2: e245.
- HALWARD, T., T. STALKER, E. LARUE, AND G. KOCHERT. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Molecular Biology* 18: 315-325.
- HAMMER, K. 2003. A paradigm shift in the discipline of plant genetic resources. *Genetic resources and crop evolution* 50: 3–10.
- HAMMONS, R.O. 1970. Registration of Spancross peanuts. *Crop Science* 10: 459.
- HE, G., R. MENG, M. NEWMAN, G. GAO, R. PITTMAN, AND C. PRAKASH. 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). *BMC Plant Biol* 3: 3.
- HE, G., AND C.S. PRAKASH. 1997. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). *Euphytica* 97: 143 - 149.
- HE, G., AND C. PRAKASH. 2001. Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. *Genetic Resources and Crop Evolution* 48: 347-352.
- HOLBROOK, C.C., AND J.P. NOE. 1992. Resistance to the Peanut Root-Knot Nematode (*Meloidogyne arenaria*) in *Arachis hypogaea* 1. *Peanut Science* 19: 35–37.
- HOLBROOK, C.C., J.P. NOE, M.G. STEPHENSON, AND W.F. ANDERSON. 1996. Identification and Evaluation of Additional Sources of Resistance to Peanut Root-knot Nematode in *Arachis hypogaea* L. 1. *Peanut Science* 23: 91–94.
- HOLBROOK, C., AND H.T. STALKER. 2003. Peanut Breeding and Genetic Resources. *Plant breeding reviews* 22: 297-356.
- HONG, Y., X. CHEN, X. LIANG, H. LIU, G. ZHOU, S. LI, S. WEN, C.C. HOLBROOK, AND B. GUO. 2010. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. *BMC Plant Biology* 10: 17.
- HOPKINS, M.S., A.M. CASA, T. WANG, S.E. MITCHELL, R.E. DEAN, G.D. KOCHERT, AND S. KRESOVICH. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. *Crop Science* 39: 1243-1247.
- HOUGAARD, B.K., L.H. MADSEN, N. SANDAL, M. DE CARVALHO MORETZSOHN, J. FREDSLUND, L. SCHAUSER, A.M. NIELSEN, T. ROHDE, S. SATO, S. TABATA, D.J. BERTIOLI, AND J. STOUGAARD. 2008. Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and *Arachis*. *Genetics* 179: 2299-2312.

- HU, F.B., M.J. STAMPFER, J.A. MANSON, E.B. RIMM, G.A. COLDITZ, B.A. ROSNER, F.E. SPEIZER, C.H. HENNEKENS, AND W.C. WILLETT. 1998. Frequent nut consumption and risk of coronary heart disease in women: prospective cohort study. *Bmj* 317: 1341.
- HUANG, X.Q., H. CÖSTER, M.W. GANAL, AND M.S. RÖDER. 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *TAG Theoretical and Applied Genetics* 106: 1379–1389.
- HUSTED, L. 1936. Cytological studies of the peanut *Arachis*. II. Chromosome number, morphology, and behavior and their application to the origin of cultivated forms. *Cytologia* 7: 396–423.
- ISLEIB, T.G., H.E. PATTEE, AND F.G. GIESBRECHT. 1995. Ancestral Contributions to Roasted Peanut Attribute 1. *Peanut Science* 22: 42–48.
- ISLEIB, T.G., H.E. PATTEE, D.W. GORBET, AND F.G. GIESBRECHT. 2000. Genotypic Variation in Roasted Peanut Flavor Quality Across 60 Years of Breeding 1. *Peanut Science* 27: 92–98.
- JARVIS, A., M.E. FERGUSON, D.E. WILLIAMS, L. GUARINO, P.G. JONES, H.T. STALKER, J.F.M. VALLS, R.N. PITTMAN, C.E. SIMPSON, AND P. BRAMEL. 2003. Biogeography of Wild *Arachis*: Assessing Conservation Status and Setting Future Priorities. *Crop Science* 43: 1100–1108.
- JEUKEN, M.J., AND P. LINDHOUT. 2004. The development of lettuce backcross inbred lines (BILs) for exploitation of the *Lactuca saligna* (wild lettuce) germplasm. *TAG Theoretical and Applied Genetics* 109: 394–401.
- KNAUFT, D.A., AND J.C. WYNNE. 1995. Peanut breeding and genetics. *Advances in agronomy* 55: 393–445.
- KNAUFT, D.A., K.M. MOORE, AND D.W. GORBET. 1993. Further Studies On The Inheritance Of Fatty Acid Composition In Peanut 1. *Peanut Science* 20: 74–76.
- KOCHERT, G., H.T. STALKER, M. GIMENES, M.L. GALGARO, C.R. LOPES, AND K. MOORE. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *American journal of botany* 83: 1282–1291.
- KOCHERT, G., T. HALWARD, W.D. BRANCH, AND C.E. SIMPSON. 1991. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. *TAG Theoretical and Applied Genetics* 81: 565–570.
- KOUMPROGLOU, R., T.M. WILKES, P. TOWNSON, X.Y. WANG, J. BEYNON, H.S. POONI, H.J. NEWBURY, AND M.J. KEARSEY. 2002. STAIRS: a new genetic resource for functional genomic studies of *Arabidopsis*. *The Plant Journal* 31: 355–364.
- KOVACH, M., AND S. MCCOUCH. 2008. Leveraging natural diversity: back through the

- bottleneck. *Current Opinion in Plant Biology* 11: 193-200.
- KRAPOVICKAS, A., AND W. GREGORY. 1994. Taxonomía del género *Arachis* (Leguminosae). *Bonplandia* 8: 1-186.
- KRAPOVICKAS, A. 1968. Origen, variabilidad y difusión del mani (*Arachis hypogaea*). In CONGRESSO INTERNACIONAL DE AMERICANISTAS, 517–534.
- KRISHNAMURTHY, L., V. VADEZ, M.J. DEVI, R. SERRAJ, S. NIGAM, M. SHESHSAYEE, S. CHANDRA, AND R. ARUNA. 2007. Variation in transpiration efficiency and its related traits in a groundnut (*Arachis hypogaea* L.) mapping population. *Field Crops Research* 103: 189-197.
- LAVIA, G.I. 1998. Karyotypes of *Arachis palustris* and *A. praecox* (section *Arachis*), two species with basic chromosome number $x=9$. *Cytologia* 63: 177–181.
- LEAL-BERTIOLI, S., A.C. JOSE, D. ALVES-FREITAS, M. MORETZSOHN, P. GUIMARAES, S. NIELEN, B. VIDIGAL, R. PEREIRA, J. PIKE, A. FAVERO, M. PARNISKE, R. VARSHNEY, AND D. BERTIOLI. 2009. Identification of candidate genome regions controlling disease resistance in *Arachis*. *BMC Plant Biology* 9: 112.
- LI, J.Z., X.Q. HUANG, F. HEINRICHS, M.W. GANAL, AND M.S. RÖDER. 2005. Analysis of QTLs for yield, yield components, and malting quality in a BC3-DH population of spring barley. *TAG Theoretical and Applied Genetics* 110: 356-363.
- LIU, S., R. ZHOU, Y. DONG, P. LI, AND J. JIA. 2006. Development, utilization of introgression lines using a synthetic wheat as donor. *TAG Theoretical and Applied Genetics* 112: 1360-1373.
- LIU, Y., A. GUR, G. RONEN, M. CAUSSE, R. DAMIDAUX, M. BURET, J. HIRSCHBERG, AND D. ZAMIR. 2003. There is more to tomato fruit colour than candidate carotenoid genes. *Plant Biotechnology Journal* 1: 195-207.
- LOPEZ, Y., O.D. SMITH, S.A. SENSEMAN, AND W.L. ROONEY. 2001. Genetic factors influencing high oleic acid content in Spanish market-type peanut cultivars. *Crop science* 41: 51–56.
- LU, J., AND B. PICKERSGILL. 1993. Isozyme variation and species relationships in peanut and its wild relatives (*Arachis* L. Leguminosae). *TAG Theoretical and Applied Genetics* 85: 550-560.
- MADHAVA, H.B., M.S. SHESHSAYEE, A.G. SHANKAR, T.G. PRASAD, AND M. UDAYAKUMAR. 2003. Use of SPAD chlorophyll meter to assess transpiration efficiency of peanut. In Breeding of drought-resistant peanuts: proceedings of a Collaborative Review Meeting held on 25-27 February, 2002 at Hyderabad, Andhra Pradesh India, 3.
- MALLIKARJUNA, N., AND D. SASTRI. 2002. Morphological, cytological and disease resistance studies of the intersectional hybrid between *Arachis hypogaea* L. and *A. glabrata* Benth. *Euphytica* 126: 161-167.

- MEHAN, V.K., P.M. REDDY, P. SUBRAHMANYAM, D. McDONALD, AND A.K. SINGH. 1996. Identification of new sources of resistance to rust and late leaf spot in peanut. *International Journal of Pest Management* 42: 267.
- MILLA, S.R., T.G. ISLEIB, AND H.T. STALKER. 2005. Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers. *Genome* 48: 1-11.
- MONCADA, P., C.P. MARTINEZ, J. BORRERO, M. CHATEL, H. GAUCH JR, E. GUIMARAES, J. TOHME, AND S.R. MCCOUCH. 2001. Quantitative trait loci for yield and yield components in an *Oryza sativa* x *Oryza rufipogon* BC₂F₂ population evaluated in an upland environment. *TAG Theoretical and Applied Genetics* 102: 41-52.
- MONFORTE, A.J., E. FRIEDMAN, D. ZAMIR, AND S.D. TANKSLEY. 2001. Comparison of a set of allelic QTL-NILs for chromosome 4 of tomato: deductions about natural variation and implications for germplasm utilization. *TAG Theoretical and Applied Genetics* 102: 572-590.
- MONFORTE, A.J., AND S.D. TANKSLEY. 2000a. Fine mapping of a quantitative trait locus (QTL) from *Lycopersicon hirsutum* chromosome 1 affecting fruit characteristics and agronomic traits: breaking linkage among QTLs affecting different traits and dissection of heterosis for yield. *TAG Theoretical and Applied Genetics* 100: 471-479.
- MONFORTE, A.J., AND S.D. TANKSLEY. 2000b. Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. *Genome* 43: 803-813.
- MOORE, K.M., AND D.A. KNAUFT. 1989. The inheritance of high oleic acid in peanut. *Journal of Heredity* 80: 252.
- MORETZSOHN, M., M. HOPKINS, S. MITCHELL, S. KRESOVICH, J. VALLS, AND M. FERREIRA. 2004. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. *BMC Plant Biology* 4: 11.
- MORETZSOHN, M., L. LEOI, K. PROITE, P. GUIMARÃES, S. LEAL-BERTIOLI, M. GIMENES, W. MARTINS, J. VALLS, D. GRATTAPAGLIA, AND D. BERTIOLI. 2005. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *TAG Theoretical and Applied Genetics* 111: 1060-1071.
- MORETZSOHN, M., A. BARBOSA, D. ALVES-FREITAS, C. TEIXEIRA, S. LEAL-BERTIOLI, P. GUIMARAES, R. PEREIRA, C. LOPES, M. CAVALLARI, J. VALLS, D. BERTIOLI, AND M. GIMENES. 2009. A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. *BMC Plant Biology* 9: 40.
- MURTY, U., AND M. JAHNAVI. 1986. The A genome of *Arachis hypogaea* L. *Cytologia* 51: 241-250.
- NAGESWARA RAO, N., AND G. WRIGHT. 1994. Stability of the relationship between specific leaf area and carbon isotope discrimination across environments in peanut. *Crop Science* 34: 98-103.

- NAGESWARA RAO, R.C., H.S. TALWAR, AND G.C. WRIGHT. 2001. Rapid Assessment of Specific Leaf Area and Leaf Nitrogen in Peanut (*Arachis hypogaea* L.) using a Chlorophyll Meter. *Journal of Agronomy and Crop Science* 186: 175-182.
- NAGESWARA RAO, R.C., G.C. WRIGHT, AND A.L. CRUICKSHANK. 2000. Genetic enhancement of drought resistance in Australian peanuts. *In* Proceedings of American Peanut Research Education Society, 71.
- NAIDU, R.A., F.M. KIMMINS, C.M. DEOM, P. SUBRAHMANYAM, A.J. CHIYEMBEKEZA, AND P.J.A. VAN DER MERWE. 1999. Groundnut Rossette: A Virus Disease Affecting Groundnut Production in Sub-Saharan Africa. *Plant Disease* 83: 700-709.
- NARASIMHAMOORTHY, B., B. GILL, A. FRITZ, J. NELSON, AND G. BROWN-GUEDIRA. 2006. Advanced backcross QTL analysis of a hard winter wheat \times synthetic wheat population. *TAG Theoretical and Applied Genetics* 112: 787-796.
- NORDEN, A.J., D.W. GORBET, D.A. KNAUFT, AND C.T. YOUNG. 1987. Variability in Oil Quality Among Peanut Genotypes in the Florida Breeding Program 1. *Peanut science* 14: 7-11.
- PASSIOURA, J.B. 1977. Grain yield, harvest index and water use of wheat. *Journal of the Australian Institute of Agricultural Science* 43: 117-120.
- PATERSON, A.H., AND R.A. WING. 1993. Genome mapping in plants. *Current opinion in Biotechnology* 4: 142-147.
- PATTEE, H.E., T.G. ISLEIB, D.W. GORBET, F.G. GIESBRECHT, AND Z. CUI. 2001. Parent Selection in Breeding for Roasted Peanut Flavor Quality 1. *Peanut Science* 28: 51-58.
- PEÑALOZA, A.P.S., AND J.F. VALLS. 2005. Chromosome number and satellited chromosome morphology of eleven species of *Arachis* (Leguminosae). *Revista del Instituto de Botánica del Nordeste* 1: 65.
- PESTSOVA, E.G., A. BÖRNER, AND M.S. RÖDER. 2006. Development and QTL assessment of *Triticum aestivum*-*Aegilops tauschii* introgression lines. *TAG Theoretical and Applied Genetics* 112: 634-647.
- PILLEN, K., A. ZACHARIAS, AND J. LEON. 2003. Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *TAG Theoretical and Applied Genetics* 107: 340-52.
- PILLEN, K., A. ZACHARIAS, AND J. LÉON. 2004. Comparative AB-QTL analysis in barley using a single exotic donor of *Hordeum vulgare* ssp. *spontaneum*. *TAG Theoretical and Applied Genetics* 108: 1591-1601.
- PLUCKNETT, D.L., N.J.H. SMITH, J.T. WILLIAMS, AND N. MURTHI ANISHETTY. 1987. Gene Banks and the world's food Princeton University press, Princeton, NJ 247 pp. *RHIZOCTONIA BATATICOIA*.
- PROITE, K., S.C.M. LEAL-BERTIOLI, D.J. BERTIOLI, M.C. MORETZSOHN, F.R.D. SILVA, N.F.

- MARTINS, AND P.M. GUIMARÃES. 2007. ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC plant biology* 7: 7.
- RAINA, S.N., V. RANI, T. KOJIMA, Y. OGIHARA, K.P. SINGH, AND R.M. DEVARUMATH. 2001. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. *Genome* 44: 763-72.
- RAO, G.U., A. BEN CHAIM, Y. BOROVSKY, AND I. PARAN. 2003. Mapping of yield-related QTLs in pepper in an interspecific cross of *Capsicum annuum* and *C. frutescens*. *TAG Theoretical and Applied Genetics* 106: 1457–1466.
- RAO, K., AND BRAMEL. 2003. Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution* 50: 707-721.
- RAO, R.C., S. SINGH, M.V.K. SIVAKUMAR, K.L. SRIVASTAVA, AND J.H. WILLIAMS. 1985. Effect of water deficit at different growth phases of peanut. I. Yield responses. *Agronomy journal* 77: 782-786.
- REVOREDO, C.L., AND S. FLETCHER. 2002. World peanut market: an overview of the past 30 years. Georgia Agricultural Experiment Stations, College of Agricultural and Environmental Sciences, the University of Georgia.
- ROBLEDO, G., AND G. SEIJO. 2010. Species relationships among the wild B genome of *Arachis* species (section *Arachis*) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. *TAG Theoretical and Applied Genetics*. Available at: <http://www.springerlink.com/content/h883047j888m9h86/> [Accessed August 18, 2010].
- SABATE, J. 2003. Nut consumption and body weight. *The American Journal of Clinical Nutrition* 78: 647S-650.
- SANDERS, T.H., H.E. PATTEE, J.R. VERCELLOTTI, AND K.L. BETT. 1995. Advances in peanut flavor quality. *Advances in Peanut Science* 528–553.
- SCHMALENBACH, I., N. KÖRBER, AND K. PILLEN. 2008. Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. *TAG Theoretical and Applied Genetics* 117: 1093-1106.
- SCHMALENBACH, I., J. LÉON, AND K. PILLEN. 2009. Identification and verification of QTLs for agronomic traits using wild barley introgression lines. *TAG Theoretical and Applied Genetics* 118: 483-97.
- SEIJO, G., G.I. LAVIA, A. FERNANDEZ, A. KRAPOVICKAS, D.A. DUCASSE, D.J. BERTIOLI, AND E.A. MOSCONE. 2007. Genomic relationships between the cultivated peanut (*Arachis hypogaea*, *Leguminosae*) and its close relatives revealed by double GISH. *American Journal of Botany* 94: 1963-1971.
- SEIJO, J.G., G.I. LAVIA, A. FERNANDEZ, A. KRAPOVICKAS, D. DUCASSE, AND E.A. MOSCONE. 2004. Physical mapping of the 5S and 18S-25S rRNA genes by FISH as evidence that

- Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). *American Journal of Botany* 91: 1294-1303.
- SEPTININGSIH, E.M., J. PRASETIYONO, E. LUBIS, T.H. TAI, T. TJUBARYAT, S. MOELJOPAWIRO, AND S.R. MCCOUCH. 2003. Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *TAG Theoretical and Applied Genetics* 107: 1419–1432.
- SEPTININGSIH, E.M., K.R. TRIJATMIKO, S. MOELJOPAWIRO, AND S.R. MCCOUCH. 2003. Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *TAG Theoretical and Applied Genetics* 107: 1433-1441.
- SHOKES, F.M., AND A.K. CULBREATH. 1997. Early and late leaf spots. *Compendium of Peanut Diseases, 2nd Edition*. APS Press, St. Paul, MN 17–20.
- SIMPSON, C., AND J. STARR. 2001. Registration of 'COAN' Peanut. *Crop Science* 41: 918.
- SIMPSON, C., J. STARR, G. CHURCH, M. BUROW, AND A. PATERSON. 2003. Registration of 'NemaTAM' Peanut. *Crop Science* 43: 1561.
- SINGH, A.K., J. SMARTT, C.E. SIMPSON, AND S.N. RAINA. 1998. Genetic variation vis-à-vis molecular polymorphism in groundnut, *Arachis hypogaea* L. *Genetic Resources and Crop Evolution* 45: 119-126.
- SMARTT, J., M.P. GREGORY, AND W.C. GREGORY. 1978. The genomes of *Arachis hypogaea* 2. The implications in interspecific breeding. *Euphytica* 27: 677-680.
- SMARTT, J., AND H.T. STALKER. 1982. Speciation and cytogenetics in *Arachis*. *Peanut science and technology* 21–49.
- SONGSRI, P., S. JOGLOY, T. KESMALA, N. VORASOOT, C. AKKASAENG, A. PATANOTHAI, AND C.C. HOLBROOK. 2008. Heritability of Drought Resistance Traits and Correlation of Drought Resistance and Agronomic Traits in Peanut. *Crop Science* 48: 2245-2253.
- STALKER, H.T. 1991. A new species in section *Arachis* of peanuts with a D genome. *American Journal of Botany* 78: 630–637.
- STALKER, H.T., J.S. DHESI, AND D.C. PARRY. 1991. An analysis of the B genome species *Arachis batizocoi* (Fabaceae). *Plant Systematics and Evolution* 174: 159-169.
- STALKER, H.T., T.D. PHILLIPS, J.P. MURPHY, AND T.M. JONES. 1994. Variation of isozyme patterns among *Arachis* species. *TAG Theoretical and Applied Genetics* 87: 746–755.
- STALKER, H.T., AND C.E. SIMPSON. 1995. Germplasm resources in *Arachis*. In *Advances in peanut science*. Patte HE and Stalker HT, eds. Stillwater, Oklahoma, USA: American Peanut Research and Education Society. Inc.
- SUBRAHMANYAM, P., G.L. HILDEBRAND, R.A. NAIDU, L.J. REDDY, AND A.K. SINGH. 1998.

- Sources of resistance to groundnut rosette disease in global groundnut germplasm*. *Annals of Applied Biology* 132: 473–485.
- SUBRAHMANYAM, P., V.R. RAO, D. McDONALD, J.P. MOSS, AND R.W. GIBBONS. 1989. Origins of resistances to rust and late leaf spot in peanut (*Arachis hypogaea*, *Fabaceae*). *Economic Botany* 43: 444–455.
- TAKAHASHI, Y., A. SHOMURA, T. SASAKI, AND M. YANO. 2001. Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. *PNAS* 98: 7922–7927.
- TANKSLEY, S.D. 1993. Mapping polygenes. *Annual Review of Genetics* 27: 205–233.
- TANKSLEY, S.D., S. GRANDILLO, T.M. FULTON, D. ZAMIR, Y. ESHED, V. PETIARD, J. LOPEZ, AND T. BECK-BUNN. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *TAG Theoretical and Applied Genetics* 92: 213–224.
- TANKSLEY, S.D., AND J.C. NELSON. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *TAG Theoretical and Applied Genetics* 92: 191–203.
- TANKSLEY, S.D., AND S.R. MCCOUCH. 1997. Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science* 277: 1063–1066.
- THOMSON, M.J., T.H. TAI, A.M. MCCLUNG, X.H. LAI, M.E. HINGA, K.B. LOBOS, Y. XU, C.P. MARTINEZ, AND S.R. MCCOUCH. 2003. Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *TAG Theoretical and Applied Genetics* 107: 479–93.
- TIAN, F., D.J. LI, Q. FU, Z.F. ZHU, F.C. FU, X.K. WANG, AND C.Q. SUN. 2006. Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *TAG Theoretical and Applied Genetics* 112: 570–580.
- UPADHYAYA, H., B. FURMAN, S. DWIVEDI, S. UDUPA, C. GOWDA, M. BAUM, J. CROUCH, H. BUHARIWALLA, AND S. SINGH. 2006. Development of a Composite Collection for Mining Germplasm Possessing Allelic Variation for Beneficial Traits in Chickpea. *Plant Genetic Resources* 4: 13–19.
- UPADHYAYA, H.D. 2005. Variability for Drought Resistance Related Traits in the Mini Core Collection of Peanut. *Crop Science* 45: 1432–1440.
- VALLS, J.F.M., AND C.E. SIMPSON. 2005. New species of *Arachis* (*Leguminosae*) from Brazil, Paraguay and Bolivia. *Bonplandia* 14: 35–63.
- VARSHNEY, R., D. BERTIOLI, M. MORETZSOHN, V. VADEZ, L. KRISHNAMURTHY, R. ARUNA, S. NIGAM, B. MOSS, K. SEETHA, K. RAVI, G. HE, S. KNAPP, AND D. HOISINGTON. 2008.

- The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). *TAG Theoretical and Applied Genetics* 118: 729-739.
- VON KORFF, M., H. WANG, J. LEON, AND K. PILLEN. 2005. AB-QTL analysis in spring barley. I. Detection of resistance genes against powdery mildew, leaf rust and scald introgressed from wild barley. *TAG Theoretical and Applied Genetics* 111: 583–590.
- VON KORFF, M., H. WANG, J. LÉON, AND K. PILLEN. 2006. AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H. vulgare ssp. spontaneum*). *TAG Theoretical and Applied Genetics* 112: 1221–1231.
- VON KORFF, M., H. WANG, J. LÉON, AND K. PILLEN. 2004. Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare ssp. spontaneum*) as donor. *TAG Theoretical and Applied Genetics* 109: 1736-1745.
- WALIYAR, F., B.B. SHEW, R. SIDAHMED, AND M.K. BEUTE. 1995. Effects of Host Resistance on Germination of *Cercospora arachidicola* on Peanut Leaf Surfaces 1. *Peanut Science* 22: 154–157.
- WAN, X.Y., J.M. WAN, C.C. SU, C.M. WANG, W.B. SHEN, J.M. LI, H.L. WANG, L. JIANG, S.J. LIU, L. M. CHEN, H. YASUI, AND A. YOSHIMURA. 2004. QTL detection for eating quality of cooked rice in a population of chromosome segment substitution lines. *TAG Theoretical and Applied Genetics* 110: 71-79.
- WOJCIECHOWSKI, M.F., M. LAVIN, AND M.J. SANDERSON. 2004. A phylogeny of legumes (*Leguminosae*) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. *American Journal of Botany* 91: 1846.
- WRIGHT, G.C., AND N.C. RACHAPUTI. 2003. Transpiration Efficiency. *Water and Plants (Biology) for the Encyclopedia of Water Science; Marcel Dekker, Inc* 982–988.
- WRIGHT, G.C., K.T. HUBICK, AND G.D. FARQUHAR. 1991. Physiological analysis of peanut cultivar response to timing and duration of drought stress. *Australian Journal of Agricultural Research* 42: 453–470.
- WRIGHT, G.C., R. NAGESWARA, AND G.D. FARQUHAR. 1994. Water-use efficiency and carbon isotope discrimination in peanut under water deficit conditions. *Crop Science* 34: 92–97.
- XIAO, J., S. GRANDILLO, S.N. AHN, S.R. MCCOUCH, S.D. TANKSLEY, J. LI, AND L. YUAN. 1996. Genes from wild rice improve yield. *Nature* 384: 223–224.
- XIAO, J., J. LI, S. GRANDILLO, S.N. AHN, L. YUAN, S.D. TANKSLEY, AND S.R. MCCOUCH. 1998. Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150: 899.
- YANG, Z. 2010. Are peanut allergies a concern for using peanut-based formulated foods in developing countries? *Food and Nutrition Bulletin* 31: S147-153.

- YANO, M. 2001. Genetic and molecular dissection of naturally occurring variation. *Current Opinion in Plant Biology* 4: 130-135.
- YANO, M., S. KOJIMA, Y. TAKAHASHI, H. LIN, AND T. SASAKI. 2001. Genetic Control of Flowering Time in Rice, a Short-Day Plant. *Plant Physiology* 127: 1425-1429.
- YATES, H.E., A. FRARY, S. DOGANLAR, A. FRAMPTON, N.T. EANNETTA, J. UHLIG, AND S.D. TANKSLEY. 2004. Comparative fine mapping of fruit quality QTLs on chromosome 4 introgressions derived from two wild tomato species. *Euphytica* 135: 283–296.
- ZAMIR, D. 2001. Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics* 2: 983-989.

RESUME

L'arachide (*Arachis hypogaea* L.) est une plante allotétraploïde ($2n = 4x = 40$) issue d'un événement récent d'hybridation entre deux espèces sauvages suivi d'un doublement spontané des chromosomes. La variabilité génétique existant dans le compartiment cultivé est faible. Les espèces diploïdes sauvages apparentées à l'arachide cultivée représentent un important réservoir d'allèles nouveaux utilisables pour élargir le pool génique du cultigène. L'objectif général de cette thèse est d'augmenter les marges de progression en amélioration de l'arachide par recours aux ressources génétiques sauvages.

Nous avons développé une population BC_1F_1 à partir du croisement entre Fleur11 et un amphidiploïde synthétique (*A. ipaensis* x *A. duranensis*)^{x4} qui associe les génomes des plus probables progéniteurs sauvages de l'espèce cultivée. Nous avons d'abord construit une carte génétique comprenant 298 loci positionnés sur 21 groupes de liaison avec une taille totale de 1843.7 cM. Nous avons ensuite conduit une analyse AB-QTL pour plusieurs caractères impliqués dans la productivité, l'adaptation et la domestication de l'arachide. Au total, nous avons cartographié 95 QTL. Pour plusieurs QTL, les effets positifs sont associés aux allèles provenant des espèces sauvages. Nous avons aussi identifié trois régions du génome qui portent des empreintes de la domestication. Nous avons enfin développé une population de 122 lignées d'introgression à l'aide d'une stratégie de sélection assistée par marqueurs. L'ensemble des groupes de liaison sont couverts avec des fragments chevauchants, issus des donneurs sauvages, d'une taille moyenne de 39.2 cM et chaque lignée comprend en moyenne 1.2 fragments.

Nous avons par ailleurs discuté l'utilisation de cette collection de lignées d'introgression pour des applications de sélection et de génétique. Nos résultats ouvrent de nouvelles perspectives pour l'amélioration de l'arachide par croisement avec les espèces sauvages apparentées.

Mots-clés : *Arachis hypogaea*, cartographie de QTL, domestication, ressources génétiques sauvages, polyploïdes, tolérance à la sécheresse, amélioration.

ABSTRACT

Peanut (*Arachis hypogaea* L.) is considered to be an allotetraploid ($2n = 4x = 40$) originated from a single hybridization event between two wild diploids followed by spontaneous chromosomes duplication. Cultivated peanut harbours a limited genetic diversity. Peanut wild relatives represent an important source of novel alleles that could be used to broaden the gene pool of the cultigen. The general objective of this work is to enhance the rate of progress in peanut breeding using wild species resources.

We developed a BC_1F_1 population from the cross between Fleur11 and a synthetic amphidiploid (*A. ipaensis* x *A. duranensis*)^{x4} that combines the genomes of the most probable wild progenitors of the cultigen. We first developed a SSR based genetic map of 298 loci on 21 linkage groups, spanning a total map distance of 1843.7 cM. We then conducted a detailed AB-QTL analysis for several traits involved in peanut productivity and adaptation as well as in the domestication syndrome. We mapped a total of 95 QTLs. About half of the QTL positive effects were associated with alleles of the wild parents. We identified three genomic regions which bear footprints of domestication. We finally developed, through a marker assisted backcross strategy, an exotic introgression library of 122 lines. This population, which has in average 1.2 fragments per line, allows covering all linkage groups with overlapping wild donor fragments of 39.2 cM length in average.

The utilization of the exotic introgression lines library for breeding and genetics is discussed. Our findings open new avenues for peanut improvement using wild relatives.

Keywords: *Arachis hypogaea*, QTL mapping, domestication, wild relatives, polyploid, drought tolerance, breeding.